Univerzita Karlova v Praze

Matematicko-fyzikální fakulta

BAKALÁRSKÁ PRÁCA

Martin Onderko

Testy významnosti parametrů ARMA modelů založené na bayesovském přístupu

Katedra pravděpodobnosti a matematické statistiky

Vedúci bakalárskej práce: Mgr. Krtek Jiří

Študijný program: Matematika

Študijný obor: finančná matematika

Praha 2012

Poďakovanie

Dovoľujem si na tomto mieste poďakovať Mgr. Jiřímu Krtkovi, vedúcemu bakalárskej práce, za jeho pomoc i užitočné rady pri vzniku tejto bakalárskej práce.

Čestné vyhlásenie

Čestne vyhlasujem, že som túto bakalársku prácu vypracoval samostatne a výhradne s použitím citovaných prameňov, literatúry a ďalších odborných zdrojov.

Beriem na vedomie, že sa na moju prácu vzťahujú práva a povinnosti vyplývajúce zo zákona č. 121/2000 Sb., autorského zákona v platnom znení, hlavne skutočnosť, že Univerzita Karlova v Praze má právo na uzatvorenie licenčnej zmluvy o použití tejto práce ako školského diela podľa § 60 odst. 1 autorského zákona.

V Prahe dňa 24.05.2012 ....

Názov práce: Testy významnosti parametrů ARMA modelů založené na bayesovském přístupu
Autor: Martin Onderko
Katedra: Katedra pravděpodobnosti a matematické statistiky
Vedúci bakalárskej práce: Mgr. Jiří Krtek, KPMS-externý pracovník
Abstrakt: Táto bakalárska práca sa zameriava na bayesovskú analýzu a na jej využitie v oblasti pravdepodobnosti a štatistiky. Okrajovo sa dotýka náhodných procesov, bližšie charakterizuje model ARMA a venuje sa problematike odhadov jeho parametrov bayesovským prístupom. Získané vedomosti a odvodené charakteristiky následne uplatňuje v testoch významnosti parametrov. Nepochybne sa tak dotýka oblasti testovania hypotéz a slúži hlavne ako nástroj k presnejšiemu určeniu modelu ARMA. Táto práca by sa mala uplatniť všade tam, kde je charakterizovanie štatistickej významnosti parametrov modelu ARMA nevyhnutnosťou.
Kľúčové slová: model ARMA, bayesovská analýza, odhady parametrov, testy významnosti parametrov

Title: Tests of significance of ARMA models parameters based on Bayesian approach
Author: Martin Onderko
Department: Department of Probability and Mathematical Statistics
Supervisor: Mgr. Jiří Krtek, DPMS-external specialist
Abstract: This thesis is focused on Bayesian analysis and its use in probability and statistics. It also marginally discusses random processes, furtherly describes ARMA model and defines the issue of estimation of the parameters of Bayesian approach. Acquired knowledge and derived characteristics subsequently applies in testing of significance of parameters. Thus it undoubtably affects the area of hypothesis testing and serves mainly as a tool to determine the ARMA model more accurately. This work should be regulary applied when detecting the necessity of testing of statistical significance of parameters of ARMA model.
Keywords: ARMA model, Bayesian analysis, estimates of the parameters, tests of significance of parameters

OBSAH

1 ZÁKLADNÉ POJMY 2

1.1 Úvod do teórie pravdepodobnosti 2

1.2 Prehľad pravdepodobnostných rozdelení 6

2 TEÓRIA NÁHODNÝCH PROCESOV 10

2.1 Náhodný proces 10

2.1.1 Stacionárny proces a funkcie procesu 10

2.1.2 Model ARMA 11

3 BAYESOVSKÁ ANALÝZA 15

3.1 Objektívna a subjektívna apriórna informácia 15

3.2 Bayesova veta a jej význam 16

3.2.1 Bayesova veta 17

3.2.2 Význam Bayesovej vety 18

3.3 Konjugované systémy apriórnych hustôt 18

3.3.1 Konštrukcia konjugovaných systémov apriórnych hustôt 19

3.3.2 Prehľad konjugovaných systémov apriórnych hustôt 22

3.4 Princíp neurčitosti 24

3.5 Jeffreysova apriórna hustota 27

3.5.1 Jeffreysova veta 28

4 ODHADY PARAMETROV BAYESOVSKÝM PRÍSTUPOM 30

4.1 Odhady parametrov AR modelu 30

4.2 Odhady parametrov MA a ARMA modelu 35

5 TESTY VÝZNAMNOSTI PARAMETROV 38

5.1 Úvod do teórie testovania 38

5.2 Nástroje k testovaniu významnosti parametrov 39

5.3 Uvedenie testov významnosti 43

6 ZÁVER 46

LITERATÚRA 47

ÚVOD

Cieľom tejto bakalárskej práce je predstaviť bayesovský pohľad na to, ako testovať významnosť parametrov modelu ARMA. Model má veľké využitie v oblasti modelovania finančných a ekonomických časových radov. Práve preto je dôležitý jeho správny odhad a to pomocou určenia významnosti jednotlivých parametrov, ktoré v modeli figurujú. Keďže sa práca opiera o bayesovskú analýzu, venuje sa do značnej miery jej výhodám, ale samozrejme aj rôznym úskaliam.

Prvá kapitola práce predstavuje zhrnutie základných definícií a tvrdení, bez ktorých nie je možné zavádzať dôležité a kľúčové pojmy v jej ďalších častiach. Zaoberá sa teóriou a základnými rozdeleniami pravdepodobnosti, ktorých uplatnenie si nájde opodstatnenie takmer vo všetkých kapitolách.

Nasledujúca kapitola v úvode stručne opisuje teóriu náhodných procesov a postupne prechádza až k definovaniu modelu ARMA. Venuje sa jeho dôležitým charakteristikám ako aj jednotlivým špecifickým modelom, ktoré ho spoločne vytvárajú.

V tretej kapitole sa zavádza bayesovská analýza. Rozsiahlosť a významnosť jej osobitého pravdepodobnostného myslenia zaberá značný priestor, ktorý je pre ňu v práci vymedzený. Kapitola sa zaoberá využitím mimoriadne dôležitej Bayesovej vety a postupne odhaľuje rôzne spôsoby voľby počiatočných informácií o parametroch, ktoré môžu vystupovať v modeloch a ako sa následne o nich vyvodzujú závery.

Predposledná v poradí štvrtá kapitola je zameraná na odhady parametrov modelov, na ktoré sa práca sústreďuje. Pre AR model je pomocou kľúčových viet vysvetlená cesta, ako sa dá dopracovať k veličinám, ktoré z bayesovského pohľadu vystupujú ako odhady parametrov tohto modelu. Pre MA a ARMA model je ukázaná problematickosť takéhoto spôsobu a predložená iná a to numerická metóda.

V záverečnej kapitole sú predostreté spôsoby a charakteristiky, pomocou ktorých sa testuje významnosť parametrov modelu AR s ohľadom na bayesovský prístup. Takisto sú v nej zhrnuté základy štatistiky, konkrétne z oblasti testovania hypotéz, ktoré sa následne využívajú pri uvedení testov významnosti.

K väčšine matematických viet sú priložené dôkazy. Pri ich dokazovaní sme sa pridržiavali zdrojov uvedených v použitej literatúre.

1 Základné pojmy

V prvej časti tejto kapitoly si rozoberieme základné definície a tvrdenia z teórie pravdepodobnosti, ktoré potrebujeme pri definovaní a odvodzovaní pojmov, ktorým sa v práci venujeme. Základné pravdepodobnostné rozdelenia, na ktoré sa práca sústreďuje, sú uvedené v závere kapitoly. Podkladom pre tvorbu tejto kapitoly je prameň [1] a zdroj [5].

1.1 Úvod do teórie pravdepodobnosti

<u>Definícia 1.1.</u> ***Pravdepodobnostným priestorom*** rozumieme trojicu $(\Omega,\ \Lambda,\ P)$. Množinu Ω nazývame priestor elementárnych javov s prvkami $\omega \in \Omega$. Na priestore Ω definujeme neprázdny systém jeho podmnožín Ω s vlastnosťou, že $\varnothing \in \Omega$. Tento systém je uzavretý vzhľadom k spočetným zjednoteniam a k tvoreniu doplnkov. Takto definovaný systém podmnožín Ω sa nazýva σ -algebra. Každej podmnožine $A \in \Lambda$ je priradené číslo $P(A),$ kde

$$P(A) \geq 0 \ \text{ a } \ P(\Omega) = 1.$$

Pre $A_1,\ A_2,\ A_3\ \ldots \in \Lambda,\ A_i \cap A_j = \varnothing\ \ \forall i \neq j$ platí, že

$$P\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} P(A_i).$$

<u>Definícia 1.2.</u> ***Borelovská σ -algebra $\mathcal{B}_0^{\mathrm{n}}$*** v $\mathbb{R}^n$ je definovaná ako

$$\mathcal{B}_0^{\mathrm{n}} = \sigma \ \ (a_1,b_1)\times(a_2,b_2)\times\ldots\times(a_n,b_n);\ a_i < b_i \in \mathbb{R},\ i = 1,2,\ldots,n \ \ ,$$

kde systém $\mathcal{B}_0^{\mathrm{n}}$ je neprázdny a je uzavretý na doplnky a spočítateľné zjednotenia podľa definície σ -algebry. Prvky $\mathcal{B}_0^{\mathrm{n}}$ sa nazývajú borelovské množiny.

<u>Definícia 1.3.</u> ***Reálna náhodná veličina*** X je merateľné zobrazenie pravdepodobnostného priestoru $(\Omega,\ \Lambda,\ P)$ do priestoru $(\mathbb{R},\ \mathcal{B}),$ kde $\mathcal{B}$ je systém borelovských množín na $\mathbb{R}$. Merateľnosť znamená, že $\forall\, B\, \in \mathcal{B}$ platí

$$X^{-1}(B) = \ \ \omega \in \Omega : X(\omega) \in B \ \ \in \Lambda.$$

<u>Definícia 1.4.</u> ***Rozdelením náhodnej veličiny*** $X : (\Omega,\ \Lambda) \to (\mathbb{R},\ \mathcal{B})$ myslíme indukovanú pravdepodobnostnú mieru P_X na $(\mathbb{R},\ \mathcal{B}),$ definovanú vzťahom

$$P_X(B) = P \ \ \ \omega \in \Omega : X(\omega) \in B \ \ \ ,\ B \in \mathcal{B} \ \text{ a } \ P_X(B) = P[X \in B].$$

Definícia 1.5. Miera μ na $(\mathbb{R},\ \mathcal{B})$, je σ ***-konečná***, práve keď $\exists\ B_1,\ B_2,\ \ldots \in \mathcal{B}$, pre ktoré platí

$$\bigcup_{i=1}^{\infty} B_i = \mathbb{R} \text{ a } \mu(B_i) < \infty\ \forall i.$$

Definícia 1.6. Miera P_X je ***absolútne spojitá*** vzhľadom k miere μ na $(\mathbb{R},\ \mathcal{B})$, práve keď

$$\forall B \in \mathcal{B} : \mu(B) = 0 \Rightarrow P_X(B) = 0.$$

Tvrdenie 1.1. Nech μ je σ -konečná miera na $(\mathbb{R},\ \mathcal{B})$ a P_X je absolútne spojitá vzhľadom k μ. Potom $\exists$ nezáporná reálna merateľná funkcia $f_X(x)$ taká, že pre všetky merateľné funkcie $h : (\mathbb{R},\ \mathcal{B}) \to (\mathbb{R},\ \mathcal{B}_0)$ platí

$$\int_{\mathbb{R}} h(x) dP_X(x) = \int_{\mathbb{R}} h(x) f_X(x) d\mu(x).$$

Definícia 1.7. Funkcia $f_X(x)$ sa nazýva ***hustota náhodnej veličiny*** X a je určená jednoznačne μ -skoro všade.

Poznámka 1.1. Pokiaľ je miera μ čítacou mierou, ktorá je na najväčšej spočetnej množine $S = \ x_1,\ x_2,\ \ldots$ definovaná ako

$$\mu_S(B) = \#\ \ i : x_i \in B\ \ ,$$

potom má náhodná veličina X ***diskrétne rozdelenie*** a v prípade, že mierou μ je Lebesguoeova miera λ, má náhodná veličina ***spojité rozdelenie***.

Definícia 1.8. ***Modus náhodnej veličiny*** X sa označuje ako $\hat{x}$ a u diskrétneho rozdelenia predstavuje bod, pre ktorý platí

$$P(X = \hat{x}) \geq P(X = x_i)\ i = 1,\ 2,\ \ldots,$$

kde x_i sú body, v ktorých je rozdelenie sústredené. V prípade, že je rozdelenie spojité, táto charakteristika má takú hodnotu $\hat{x}$, ktorá splňuje

$$f(\hat{x}) \geq f(x),\ \ -\infty < x < \infty.$$

Modus nemusí byť určený jednoznačne.

Definícia 1.9. ***Strednou hodnotou*** EX náhodnej veličiny X rozumieme

$$EX = \int_{\Omega} X(\omega) dP(\omega),$$

pokiaľ je integrál konečný.

Tvrdenie 1.2. Jednou z vlastností strednej hodnoty sú vzťahy

$$E(a+bX)=a+b(EX)\ \forall a,\ b\in\mathbb{R} \text{ a } E(X+Y)=EX+EY, \tag{1.1}$$

ak sú splnené podmienky $E|X|<\infty$ a $E|Y|<\infty$.

Definícia 1.10. Výraz $E(X-EX)^2$ definuje ***rozptyl náhodnej veličiny*** X a zvykne sa označovať ako

$$\operatorname{var} X=\sigma_X^2=E(X-EX)^2.$$

Definícia 1.11. ***Smerodajná odchýlka*** je daná ako druhá odmocnina z rozptylu s označením σ_X.

Definícia 1.12. ***Kovarianciou náhodných veličín*** X a Y rozumieme

$$\operatorname{cov}(X,Y)=E(X-EX)(Y-EY)$$

a ***koreláciou*** týchto náhodných veličín myslíme podiel

$$\operatorname{cor}(X,Y)=\frac{E(X-EX)(Y-EY)}{\sigma_X\sigma_Y}.$$

σ_X a σ_Y sú kladné a konečné smerodajné odchýlky náhodných veličín X a Y. V prípade, že korelácia dvoch náhodných veličín je nulová, tak tieto náhodné veličiny sú nekorelované veličiny.

Definícia 1.13. Funkciu $F_X:\mathbb{R}\to\mathbb{R}$, ktorá je definovaná vzťahom

$$F_X(x)=P(X\le x),$$

nazývame ***distribučná funkcia*** náhodnej veličiny X. Strednú hodnotu náhodnej veličiny X môžeme takisto vyjadriť ako

$$EX=\int_{\mathbb{R}} x\,dF(x).$$

Definícia 1.14. ***Kvantilovou funkciou*** F_X^{-1} rozumieme funkciu danú predpisom

$$F_X^{-1}(u)=\inf\{x:F(x)\ge u\},\quad 0<u<1.$$

Definícia 1.15. V prípade, že v definícii reálnej náhodnej veličiny X rozšírime priestor $(\mathbb{R},\mathcal{B})$, na priestor $(\mathbb{R}^n,\mathcal{B}^n)$, s dimenziou $n\in\mathbb{N}$, hovoríme o ***reálnom náhodnom vektore*** $\mathbf{X}=(X_1,\ ...,\ X_n)^{\mathrm{T}}$. Hustote a distribučnej funkcii tohto náhodného vektoru $\mathbf{X}$ sa hovorí združená hustota a združená distribučná funkcia. Hustoty a distribučné funkcie jednotlivých náhodných veličín X_i, ktoré sú zložkami náhodného vektoru $\mathbf{X}$, sa označujú ako marginálne.

Definícia 1.16. Náhodné veličiny $X_1,\ ...,\ X_n$ sú ***vzájomne nezávislé,*** práve keď

$\forall \mathbf{x} = (x_1, \ ..., \ x_n) \in \mathbb{R}^n$ platí, že

$$F_{\mathbf{X}} \ \mathbf{x} \ = F_{X_1}(x_1) \cdot ... \cdot F_{X_n}(x_n),$$

kde $F_{\mathbf{X}} : \mathbb{R}^n \to \mathbb{R}$ a $F_{\mathbf{X}}(x) = P(X_1 \le x_1, \ ..., \ X_n \le x_n)$ je práve ona distribučná funkcia náhodného vektoru $\mathbf{X}$. Podobne náhodné vektory $\mathbf{X}_1$, $\mathbf{X}_2$ so zložkami n_1 a n_2 sú nezávislé, práve keď $\forall \mathbf{x} = (x_1, \ ..., \ x_{n_1+n_2}) \in \mathbb{R}^{n_1+n_2}$ platí

$$F_{\mathbf{X}}(\mathbf{x}) = F_{\mathbf{X}_1}(\mathbf{x}_1) F_{\mathbf{X}_2}(\mathbf{x}_2), \ \mathbf{X} = (\mathbf{X}_1^{\mathrm{T}}, \mathbf{X}_2^{\mathrm{T}})^{\mathrm{T}}, \mathbf{x} = (\mathbf{x}_1^{\mathrm{T}}, \mathbf{x}_2^{\mathrm{T}})^{\mathrm{T}}.$$

Tvrdenie 1.3. Náhodné veličiny $X_1, \ ..., \ X_n$ vzájomne nezávislé, práve keď vektor $\mathbf{X}$ má hustotu $f_{\mathbf{X}}$ vzhľadom k súčinovej miere $\mu = \mu_1 \otimes ... \otimes \mu_n$, kde μ_i je σ -konečná miera vzhľadom ku ktorej má náhodná veličina X_i hustotu f_{X_i} a platí

$$f_{\mathbf{X}}(\mathbf{x}) = f_{X_1}(x_1) \cdot ... \cdot f_{X_n}(x_n).$$

Poznámka 1.2. Pre potreby tejto práce je nutné uviesť, že kovariancia dvoch nezávislých náhodných veličín X, Y je rovná nule a v takomto prípade platí nasledujúce tvrdenie.

Tvrdenie 1.4. Nech teda náhodné veličiny X, Y sú nezávislé, EX^2 a EY^2 sú konečné, potom rozptyl náhodnej veličiny $X + Y$ je rovný

$$\operatorname{var}(X + Y) = \operatorname{var} X + 2\operatorname{cov}(X, Y) + \operatorname{var} Y = \operatorname{var} X + \operatorname{var} Y. \tag{1.2}$$

Tvrdenie 1.5. Platí

$$\operatorname{var}(a + bY) = b^2 \operatorname{var} Y, \ \ \operatorname{cov}(a + bX, c + dY) = bd \operatorname{cov}(X, Y), \ a, \ b, \ c, \ d \in \mathbb{R}. \tag{1.3}$$

Poznámka 1.3. Vzhľadom na rovnosti (1.2) a (1.3) platí pre nezávislé náhodné veličiny X, Y s vlastnosťami $EX^2 < 0$ a $EY^2 < 0$ vzťah

$$\operatorname{var}(X - Y) = \operatorname{var} X + \operatorname{var} Y. \tag{1.4}$$

Definícia 1.17. ***Náhodným výberom*** z rozdelenia s distribučnou funkciou F rozumieme postupnosť $\mathbf{X}_1, \ ..., \ \mathbf{X}_n$ rovnako rozdelených a nezávislých náhodných vektorov s distribučnou funkciou F.

Tvrdenie 1.6. Nech $\mathbf{Y} = (Y_1, \ ..., \ Y_n)^{\mathrm{T}}$ má hustotu $f_{\mathbf{Y}}$ vzhľadom k Lebesgueovej miere, nech $g : \mathbb{R}^n \to \mathbb{R}^n$ je prosté zobrazenie a pre skoro všetky $\mathbf{y} \in S_{\mathbf{Y}}$ je $\det \dfrac{\partial g(y)}{\partial \mathbf{y}}$ nenulový, kde $S_{\mathbf{Y}}$ je nosičom rozdelenia. Tak potom $\mathbf{X} = g(\mathbf{Y})$ má hustotu

$$f_{\mathbf{X}}(\mathbf{x}) = f_{\mathbf{Y}}\left(g^{-1}(\mathbf{x})\right)\left|\det\frac{\partial g^{-1}(\mathbf{x})}{\partial \mathbf{x}}\right| \text{ pre } \mathbf{x} \in g(S_{\mathbf{Y}})$$

$$f_{\mathbf{X}}(\mathbf{x}) = 0 \text{ inak}$$

vzhľadom k Lebesgueovej miere.

1.2 Prehľad pravdepodobnostných rozdelení

Uveďme teraz niekoľko základných jednorozmerných rozdelení a jedno dvojrozmerné rozdelenie a popíšme ich základné charakteristiky. Keďže pravdepodobnostné rozdelenie náhodnej veličiny X môže byť diskrétne alebo spojité, existuje skupina diskrétnych a skupina spojitých rozdelení.

Do skupiny diskrétnych rozdelení patrí **Alternatívne rozdelenie** s označením $\text{Alt}(p)$, kde parameter $p \in (0,1)$. Náhodná veličina X môže nadobúdať hodnoty 0 alebo 1 a hustota má tvar

$$P[X = j] = p^j (1-p)^{1-j},\ j \in \{0,\ 1\}.$$

Platí, že $EX = p$ a $\operatorname{var} X = p(1-p)$.

Ďalším rozdelením, ktoré patrí do tejto skupiny je **Binomické rozdelenie**. Označuje sa ako $\text{Bi}(n, p)$ s parametrami $n \in \mathbb{N}$, $p \in (0,1)$. Hustotou je

$$P[X = j] = \binom{n}{j} p^j (1-p)^{n-j},$$

kde $j \in \{0,\ 1,\ ...,\ n\}$, pretože náhodná veličina X môže nadobúdať práve tieto hodnoty. Stredná hodnota je np a rozptyl má hodnotu $np(1-p)$. Medzi práve uvedenými rozdeleniami existuje úzky vzťah, pretože Binomické rozdelenie je súčtom nezávislých alternatívnych veličín a z tohto dôvodu je $\text{Bi}(1, p)$ alternatívnym rozdelením s parametrom p.

Posledným diskrétnym rozdelením, ktoré sa v práci používa, je **Poissonovo rozdelenie** s parametrom $\lambda > 0$. Náhodná veličina s týmto rozdelením má strednú hodnotu a zároveň rozptyl o veľkosti λ a nadobúda hodnoty $\{0,\ 1,\ 2,\ ...\}$, preto pre $j = 0,\ 1,\ 2,\ ...$ je hustota definovaná ako

$$P[X = j] = \frac{\lambda^j}{j!} e^{-\lambda}.$$

Skupina spojitých pravdepodobnostných rozdelení obsahuje, okrem iných rozdelení, **Rovnomerné rozdelenie** s označením $\text{R}(a,b)$, kde $a,\ b \in \mathbb{R},\ a < b.$

Náhodná veličina X s týmto rozdelením nadobúda hodnoty v intervale (a,b). Jej strednou hodnotou je $\frac{a+b}{2}$, rozptyl $\frac{(b-a)^2}{12}$ a hustota v tvare

$$f(x)=\begin{cases} \frac{1}{(b-a)}, & x\in(a,b) \\ 0 & , x\notin(a,b) \end{cases}.$$

Ďalším spojitým rozdelením je **Exponencionálne rozdelenie** $\operatorname{Exp}(\lambda)$, $\lambda>0$. Stredná hodnota náhodnej veličiny X s týmto rozdelením je λ^{-1} a rozptyl má hodnotu λ^{-2}. Veličina X môže nadobúdať hodnoty v intervale $(0,\infty)$ a hustotou je

$$f(x)=\begin{matrix} \lambda e^{-\lambda x}, & x>0 \\ 0 & , x\le 0 \end{matrix}.$$

Teraz uvedieme **Gama rozdelenie**, ktoré sa označuje ako $\Gamma(a,p)$. Tieto kladné parametre určujú strednú hodnotu a rozptyl náhodnej veličiny s rozdelením $\Gamma(a,p)$, a to konkrétne hodnotou pa^{-1} a pa^{-2}. Rovnako ako v predošlom prípade náhodná veličina X s týmto rozdelením nadobúda hodnoty v intervale $(0,\infty)$ a hustotu poznáme vo vyjadrení

$$f(x)=\begin{cases} \frac{a^p}{\Gamma(p)}x^{p-1}e^{-ax}, & x>0 \\ 0 & , x\le 0 \end{cases},$$

kde $\Gamma(p)$ je gama funkcia.

Beta rozdelenie sa označuje ako $\mathrm{B}(\alpha,\beta)$. Parametre (α,β) sú kladné. Náhodná veličina s týmto rozdelením môže nadobúdať hodnoty v intervale $(0,1)$. Hustota tohto rozdelenia je

$$f(x)=\begin{cases} \frac{1}{B(\alpha,\beta)}x^{\alpha-1}(1-x)^{\beta-1}, & x\in(0,1) \\ 0 & , x\notin(0,1) \end{cases},$$

kde $B(\alpha,\beta)$ je beta funkcia. Stredná hodnota a rozptyl sa rovnajú hodnotám

$$EX=\frac{\alpha}{\alpha+\beta},\quad \operatorname{var}X=\frac{\alpha\beta}{(\alpha+\beta)^2(\alpha+\beta+1)}.$$

Medzi spojité rozdelenia, ktoré v práci využívame, patrí aj **Normálne rozdelenie** s označením $\mathrm{N}(\eta,\ \sigma^2)$. Parameter η je reálne číslo a σ je kladná. Pokiaľ má náhodná veličina X toto rozdelenie, tak jej stredná hodnota je η

a rozptylom je σ^2. Množina $\mathbb{R}$ tvorí množinu hodnôt, ktoré môže táto náhodná veličina nadobúdať. Hustotou je funkcia

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{\frac{-(x-\eta)^2}{2\sigma^2}}, \; x \in \mathbb{R}.$$

V prípade, že $\eta = 0$ a $\sigma^2 = 1$, hovoríme o normovanom normálnom rozdelení

χ^2 **rozdelenie** má označenie χ_r^2 s parametrom r nazývaným ako stupeň voľnosti. Stredná hodnota je rovná tomuto parametru a rozptyl jeho dvojnásobku. Hustotou je

$$f(x) = \begin{cases} \dfrac{1}{2^{r/2}\Gamma(r/2)} x^{(r-2)/2} e^{-x/2}, & x > 0 \\ 0 & , x \le 0 \end{cases}.$$

Jedná sa o špeciálny prípad Gama rozdelenia s parametrami $1/2$ a $r/2$.

Rozdelenie, ktoré sa označuje ako t_k, sa nazýva **Studentovo rozdelenie**. Parameter $k \in \mathbb{N}$ vyjadruje počet stupňov voľnosti rozdelenia. Pre $k = 1$ stredná hodnota neexistuje a inak má nulovú hodnotu. Ak je $k = 1, 2,$ tak rozptyl je nekonečný a pre $k \ge 3$ je rovný výrazu $\dfrac{k}{k-2}$. A nakoniec hustota rozdelenia je

$$f(x) = \frac{\Gamma\ (k+1)/2}{\Gamma(k/2)\sqrt{k\pi}} \left(1 + \frac{x^2}{k}\right)^{-(k+1)/2}, \; x \in \mathbb{R}.$$

Posledným jednorozmerným spojitým rozdelením, ktoré uvádzame, je **Fisherovo rozdelenie** (F-rozdelenie) s označením $\mathrm{F}_{m,n}$ a parametrami $m, n \in \mathbb{N}$, ktoré sa nazývajú stupne voľnosti. Hustota tohto rozdelenia je daná ako

$$f(x) = \begin{cases} \dfrac{(m/n)^{m/2}}{B(m/2, n/2)} x^{(m/2)-1} \left(1 + \dfrac{m}{n} x\right)^{-(m+n)/2}, & x > 0 \\ 0 & , x \le 0 \end{cases}.$$

Pre $n \ge 3$ má stredná hodnota vyjadrenie $\dfrac{n}{n-2}$ a inak neexistuje. Rozptyl je v tomto prípade

$$\frac{2n^2(m+n-2)}{m(n-2)^2(n-4)},$$

ak je splnené $n \ge 5$, v opačnom prípade je rovný nekonečnu.

Na záver tejto podkapitoly uvádzame spojité dvojrozmerné **Paretovo rozdelenie** s parametrami $r_1,\ r_2,\ a,$ pre ktoré platí, že $r_1 < r_2,\ a > 0.$ Pre dvojrozmerný náhodný vektor $\mathbf{X} = (X_1, X_2)^{\mathrm{T}}$ s týmto rozdelením platí, že

$$EX_1 = \frac{ar_1 - r_2}{a-1} \text{ a } EX_2 = \frac{ar_2 - r_1}{a-1} \text{ pre } a > 1$$

a ak $a > 2,$ tak

$$\operatorname{var} X_1 = \operatorname{var} X_2 = \frac{a(r_2 - r_1)^2}{(a-1)^2(a-2)}.$$

Hustotou dvojrozmerného Paretovho rozdelenia je funkcia

$$f(x_1, x_2) = \begin{cases} \dfrac{a(a+1)(r_2 - r_1)^a}{(x_2 - x_1)^{a+2}}, & (x_1, x_2) \in \mathbb{R}^2, x_1 < r_1, x_2 < r_2 \\ 0 & \text{, inak} \end{cases}.$$

2 Teória náhodných procesov

Kapitola 2 sa v úvode venuje náhodným procesom a ich základným charakteristikám. Následne pomocou nich sú charakterizované tie náhodné procesy, ktoré nevyhnutne potrebujeme k definovaniu modelu ARMA. Prameňmi pre napísanie tejto kapitoly boli diela [2], [3] a [7].

2.1 Náhodný proces

Definícia 2.1. Buď (Ω, Λ, P) pravdepodobnostný priestor a nech $\mathrm{T} \subset \mathbb{R}$. ***Náhodný proces*** definujeme ako systém náhodných veličín $\{X_t, t \in \mathrm{T}\}$, definovaných na (Ω, Λ, P).

Poznámka 2.1. Množina T býva najčastejšie totožná s množinou všetkých reálnych čísel $(\mathrm{T} = \mathbb{R})$ alebo s množinou celých čísel $(\mathrm{T} = \mathbb{Z})$.

Definícia 2.2. Náhodný proces $\{X_t, t \in \mathrm{T}\}$ sa nazýva ***časový rad***, pokiaľ množina T je totožná s množinou celých čísel $(\mathrm{T} = \mathbb{Z})$ alebo $\mathrm{T} \subset \mathbb{Z}$, pričom T interpretujeme ako čas. Niekedy sa časový rad označuje ako ***náhodná postupnosť***.

2.1.1 Stacionárny proces a funkcie procesu

Definícia 2.3. Pokiaľ existuje $\forall t \in \mathrm{T}$ stredná hodnota EX_t, potom ***strednou hodnotou procesu*** nazývame funkciu $\eta_t = EX_t$ definovanej na T.

Definícia 2.4. Ak pre proces $\{X_t, t \in \mathrm{T}\}$ platí, že $E|X_t|^2 < \infty \ \forall t \in \mathrm{T}$, definuje sa obecne komplexná ***autokovariančná funkcia procesu*** $\{X_t\}$ na $\mathrm{T} \times \mathrm{T}$ ako

$$R(s,t) = E(X_s - \eta_s)(\overline{X}_t - \overline{\eta}_t)$$

a ***autokoleračná funkcia procesu*** vzťahom

$$r(s,t) = \frac{R(s,t)}{\sqrt{R(s,s)}\sqrt{R(t,t)}}.$$

Definícia 2.5. Hodnota $R(t,t)$ sa nazýva ***rozptyl náhodného procesu*** v čase t.

Definícia 2.6. Ak funkcia $R(s,t)$ závisí len na rozdiele medzi parametra $(s-t)$, hovoríme, že proces X_t je ***kovariančne stacionárny***. Ak navyše $\eta_t = \eta \ \forall t \in \mathrm{T}$, proces sa nazýva ***slabo stacionárny*** (stacionárny).

Poznámka 2.2. Autokovariančná funkcia stacionárnych procesov sa definuje ako funkcia jednej premennej vzťahom $R(t) = R(t,0),\ t \in \mathrm{T}.$

Definícia 2.7. Hodnota $R(0) = \sigma^2$ predstavuje ***rozptyl stacionárneho procesu***.

Poznámka 2.3. V stacionárnom procese má autokorelačná funkcia predpis

$$r(t) = \frac{R(t)}{R(0)}.$$

2.1.2 Model ARMA

Pomocou zadefinovaných charakteristík teraz predstavíme náhodné procesy, ktoré využijeme pre uvedenie modelu ARMA na konci kapitoly.

Majme časový rad $\{Y_t\}$ nekorelovaných náhodných veličín s rovnakými rozptylmi σ^2. Nech stredná hodnota procesu $\eta_t = 0\ \forall t \in \mathrm{T}$ a proces je kovariančne stacionárny, kde $R(0) = \sigma^2$ a $R(t) = 0$ pre $t \neq 0.$ Časový rad s takýmito vlastnosťami sa označuje termínom ***biely šum*** (white noise) s označením $\mathrm{WN}(\sigma^2)$ a je zrejmé, že je stacionárny. Je vidieť, že autokorelačná funkcia bieleho šumu $r(t)$ nadobúda hodnotu 1 ak $t = 0$ a v prípade, že $t \neq 0$, platí $r(t) = 0.$

Uvažujme $c_0,\ c_1,\ c_2,\ ...,\ c_m$ reálne alebo komplexné čísla, pre ktoré platí, že $c_0 \neq 0$ a $c_m \neq 0.$ Vytvorme proces X_t daný vzťahom

$$X_t = \sum_{j=0}^{m} c_j Y_{t-j},\ t \in \mathrm{T}{=}\mathbb{Z}, \tag{2.1}$$

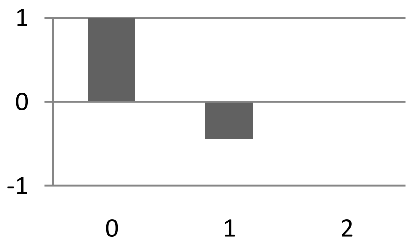
kde $Y_t \sim \mathrm{WN}(1).$ Takto zadaný proces nazývame ***postupnosť kĺzavých súčtov radu m - MA(m)***. Využitím vzťahu (1.1) získavame, že $EX_t = 0,$ a tak je stredná hodnota procesu $\eta_t = 0\ \forall t \in \mathrm{T}.$ Navyše $\operatorname{cov}(X_{s+t}, X_s)$ nezávisí na $s,$ ale len na $t,$ a preto je proces podľa definície 2.6. stacionárny. Pre $0 \le t \le m$ je autokovariančná funkcia procesu rovná

$$R(t) = \sigma^2 \sum_{j=0}^{m-t} c_{t+j}\overline{c}_j.$$

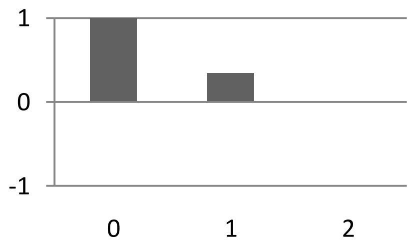
Funkcia nadobúda nulové hodnoty, ak $|t| > m$ a v prípade, že $t < 0$ na vypočítanie $R(t)$ využijeme vzťah

$$R(-t) = \overline{R(t)}.$$

Nasledujúce obrázky vyjadrujú hodnoty autokorelačnej funkcie $r(t)$ v čase t pre model MA(1) s konkrétnymi reálnymi parametrami.



Obrázok 2.1: $X_t = Y_t - 0{,}6Y_{t-1}$



Obrázok 2.2: $X_t = Y_t + 0{,}4Y_{t-1}$

Špeciálnym prípadom postupnosti kĺzavých súčtov radu m je ***postupnosť kĺzavých priemerov radu m***. Uvažované parametre $c_0, c_1, c_2, ..., c_m$ majú konštantnú hodnotu $\frac{1}{m+1}$. Kĺzavé priemery sú schopné sledovať trend pozorovaných dát.

Nech $\{Y_t\}$ je biely šum a nech pre koeficienty $a_0, a_1, a_2, ...$ platí

$$\sum_{k=0}^{\infty} |a_k|^2 < \infty.$$

Časový rad $\{X_t\}$ zadaný ako

$$X_t = \sum_{k=0}^{\infty} a_k Y_{t-k},\ t = ...,\ \text{-}1,\ 0,\ 1,\ ... \tag{2.2}$$

nazývame ***lineárny proces***. Autokovariančná funkcia tohto procesu je

$$R(t) = \sigma^2 \sum_{j=0}^{\infty} a_{t+j}\bar{a}_j,\ t = 0,\ 1,\ ...\ \text{a}\ R(-t) = \overline{R(t)},\ t < 0.$$

Využitím vzťahu (1.1) je vidieť, že stredná hodnota procesu $\eta_t = 0\ \forall t.$ Celkovo tak proces spĺňa definíciu 2.6., z toho dôvodu je stacionárny.

Uvažujme reálne čísla $d_0, d_1, d_2, ..., d_n$, kde $d_0 > 0,\ d_n \neq 0$ a $Y_t \sim \mathrm{WN}(1)$. Náhodná postupnosť zadaná tvarom

$$d_0 X_t + d_1 X_{t-1} + ... + d_n X_{t-n} = Y_t, \tag{2.3}$$

sa nazýva ***autoregresná postupnosť radu n-AR(n)***. Nech polynóm

$$d(z) = \sum_{j=0}^{n} d_j z^{n-j}$$

má všetky korene vo vnútri jednotkového kruhu. V takomto prípade sa dá dokázať (viď. prameň [2], s. 57), že existuje lineárny proces $\{X_t,\ t \in \mathbb{Z}\}$ splňujúci 2.3 a je zrejmé, že je potom postupnosť stacionárna. Keďže lineárny proces má nulovú strednú hodnotu, tak s ohliadnutím na to, že Y_t sú nekorelované náhodné veličiny, platí $EX_sY_t = 0$ pre $s < t.$ Vynásobením predpisu (2.3) autoregresnej postupnosti náhodnou veličinou X_{t-k} a uplatnením strednej hodnoty na obe strany rovnice získavame

$$d_0R(k) + d_1R(k-1) + \ldots + d_nR(k-n) = EY_tX_{t-k}. \qquad (2.4)$$

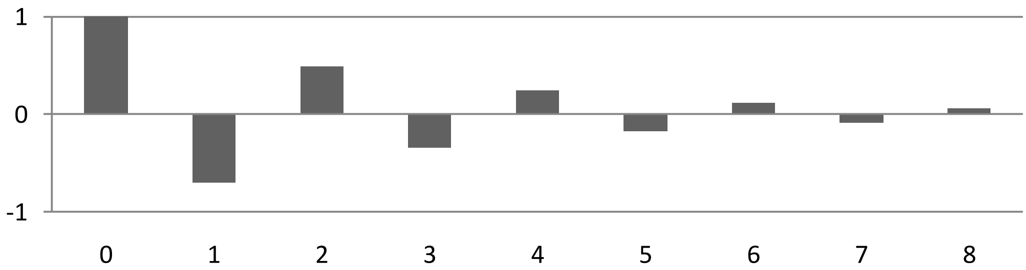
Po vydelení rovnice rozptylom procesu $R(0)$ získavame ***Yuleove-Walkerove rovnice***

$$\sum_{j=0}^{n} d_j r(k-j) = 1,\ 2,\ 3,\ \ldots\ ,$$

z ktorých sa spočíta autokoleračná funkcia $r \;\; k$. Rozptyl stacionárneho procesu $R(0)$ získame nasledovne. Predpis (2.3) vynásobíme veličinou Y_t a aplikujeme strednú hodnotu na obe strany rovnice. Nutne získame rovnosť $EX_tY_t = \sigma^2.$ Zvoľme $k = 0$ v rovnici (2.4) a položme v nej $R(t) = r(t)R(0).$ Získavame tak rovnosť

$$R(0) = \frac{\sigma^2}{d_0 + d_1r(1) + \ldots + d_nr(n)}.$$

Ukážme si teraz hodnoty autokorelačnej funkcie procesu pre konkrétny model AR(1).



Obrázok 2.3: X_t = -0,7X_{t-1} + Y_t

Nech $d_0,\ d_1,\ d_2,\ \ldots,\ d_n,$ sú reálne konštanty, kde $d_0 > 0,\ d_n \neq 0$ a nech konštanty $c_0,\ c_1,\ c_2,\ \ldots,\ c_m$ sú takisto reálne, $c_0 \neq 0,\ \ c_m \neq 0.$ Uvažujme $Y_t \sim \mathrm{WN}(1).$ Náhodná postupnosť $\{X_t, t \in \mathbb{Z}\}$ sa riadi modelom ***ARMA*(*n*,*m*)**, pokiaľ

$$d_0X_t + d_1X_{t-1} + \ldots + d_nX_{t-n} = c_0Y_t + c_1Y_{t-1} + \ldots + c_mY_{t-m}. \qquad (2.5)$$

Nech polynóm

$$d(z) = \sum_{j=0}^{n} d_j z^{n-j}$$

má všetky korene vo vnútri jednotkového kruhu a nech s polynómom

$$c(z) = \sum_{i=0}^{m} c_i z^{m-i}$$

nemá spoločné korene. Potom je možné dokázať (viď. [7], s.68), že existuje stacionárna postupnosť $\{X_t\}$ splňujúca (2.5). Je vidieť, že tento zmiešaný model je zmes autoregresie a kĺzavých súčtov. Požiadavka, aby polynómy $c(z)$ a $d(z)$ nemali spoločné korene je zásadná, pretože v opačnom prípade by došlo k zníženiu radov modelu. Autokovariančnú funkciu R je možné určiť z rovníc:

$$d_0R(k) + d_1R(k-1) + \ldots + d_nR(k-n) = \sigma^2 \sum_{j=k}^{m} c_j a_{j-k}, \; k \le m,$$
$$d_0R(k) + d_1R(k-1) + \ldots + d_nR(k-n) = 0 \qquad , k > m.$$

Celkovo sa dá povedať, že postupnosť kĺzavých súčtov a autoregresná postupnosť sú špeciálne prípady modelu ARMA.

3 Bayesovská analýza

Účelom bayesovskej analýzy je poskytnúť riešenia rôznych problémov v štatistike. Týmito problémami máme na mysli odhad parametra $\boldsymbol{\theta}$ a testovanie hypotézy o $\boldsymbol{\theta}$. Na rozdiel od klasického prístupu, Bayesov prístup spočíva v tom, že náš neznámy parameter považujeme za náhodnú veličinu. Predpokladá, že informácia o hodnote parametra, ktorú nazývame ***apriórna informácia***, je vyjadrená pravdepodobnostným rozdelením. Bayesovská analýza využíva k riešeniu samotného problému apriórnu informáciu a takisto experimentálne výsledky, ktoré sú však nezávislé na apriórnej informácii. Tieto, ale i ďalšie myšlienky a celková teória bayesovskej analýzy, ktorú v kapitole uvádzame, sú čerpané z prameňov [1], [3] a [4].

3.1 Objektívna a subjektívna apriórna informácia

Apriórna informácia môže byť objektívna, ale aj subjektívna. Pokiaľ zohľadňujeme podobné problémy z minulosti, hovoríme o objektívnej apriórnej informácii. Ak však apriórnu informáciu určuje skúsenosť, či názor subjektu, má subjektívny charakter.

Príkladom, v ktorom sa uplatňuje objektívny charakter apriórnej informácie je problém odhadu parametra θ u detí, ktorý predstavuje ich telesnú výšku v dospelosti, využitím testu s výsledkom X. Ten závisí na θ. Na základe jeho veľkého počtu realizácii je možné získať pravdepodobnostné rozdelenie X. Parameter θ môže u rôznych detí nadobúdať odlišnú hodnotu a dlhoročný výskum ukazuje, že sa dá považovať za náhodnú veličinu s normálnym rozdelením. Práve tento dlhoročný výskum indukuje objektivitu apriórnej informácie.

Uveďme teraz úlohu, odhadnúť konštantu δ, ktorá sa uplatňuje v nejakom bližšie neurčenom modeli. Majme určitú predstavu o možných hodnotách tejto konštanty a týmto hodnotám priraďme rôzne pravdepodobnosti zhody so skutočnou hodnotou konštanty, a preto ich považujme za náhodné veličiny. Náš pohľad je však subjektívny, pretože naša predstava o možných hodnotách konštanty a pravdepodobnostiach zhody so skutočnou hodnotou sa môže líšiť od pohľadu niekoho iného, kto sa zaoberá rovnakou úlohou. V tomto prípade ide o subjektívnu apriórnu informáciu o hodnote parametra.

Objektívnu, ale aj subjektívnu apriórnu informáciu by mala odrážať ***apriórna hustota*** $q(\boldsymbol{\theta})$ parametra $\boldsymbol{\theta}$. Táto hustota sa využíva pri určovaní záveroch o parametri $\boldsymbol{\theta}$ a zvykne sa pracovať s hustotami fixného funkcionálneho typu s známymi alebo neznámymi parametrami. V prípade, že nemáme žiadne informácie o $\boldsymbol{\theta}$, za apriórne rozdelenie zvolíme rovnomerné rozdelenie. Podkapitola 3.4. je venovaná práve tomuto problému.

3.2 Bayesova veta a jej význam

Definícia 3.1. Predpokladajme, že $\mathbf{U} = (U_1, \ldots, U_k)^{\mathrm{T}}$ je náhodný vektor, ktorý má hustotu $f_{\mathbf{U}}(u)$ vzhľadom k σ-konečnej miere μ_1 na $(\mathbb{R}^k, \mathcal{B}^k)$ a takisto uvažujme náhodný vektor $\mathbf{V} = (V_{k+1}, \ldots, V_n)^{\mathrm{T}}$ s hustotou $f_{\mathbf{V}}(v)$ vzhľadom k σ-konečnej miere μ_2 na $(\mathbb{R}^{n-k}, \mathcal{B}^{n-k})$, kde $\mathcal{B}^i$ predstavuje borelovskú σ-algebru v $\mathbb{R}^i$. Nech náhodný vektor $\mathbf{W} = (\mathbf{U}^{\mathrm{T}}, \mathbf{V}^{\mathrm{T}})^{\mathrm{T}}$ má hustotu $f_{\mathbf{W}}(u,v)$ vzhľadom k súčinovej miere $\mu = \mu_1 \times \mu_2$ na $(\mathbb{R}^n, \mathcal{B}^n)$. ***Podmienenou hustotou*** náhodného vektoru $\mathbf{U}$, pokiaľ je dané $\mathbf{V} = v$, nazveme ľubovoľnú merateľnú nezápornú funkciu $f(u \mid v)$, ktorá splňuje rovnosť

$$P[\mathbf{U} \in B, \mathbf{V} \in C] = \int_C \left[\int_B f(u \mid v) d\mu_1(u) \right] f_{\mathbf{V}}(v) d\mu_2(v) \;\; \forall B \in \mathcal{B}^k, \;\; \forall C \in \mathcal{B}^{n-k}.$$

Poznámka 3.1. Rovnosť nastáva, práve keď

$$f_{\mathbf{W}}(u,v) = f(u \mid v) f_{\mathbf{V}}(v) \tag{3.1}$$

μ-skoro všade, a preto podmienenú hustotu je možné vypočítať z rovnosti

$$f(u \mid v) = \frac{f_{\mathbf{W}}(u,v)}{f_{\mathbf{V}}(v)}, \tag{3.2}$$

pokiaľ v splňuje podmienku, že $f_{\mathbf{V}}(v) \neq 0$.

Podmienená hustota má nielen veľký význam pri výklade nasledujúcej ***Bayesovej vety***, ale takisto pri odhadovaní apriórnej hustoty $q(\boldsymbol{\theta})$. Niekedy sa totižto môže stať, že máme apriórnu informáciu v tvare nezávislých náhodných veličín $Y_1, \ldots, Y_n$, ktoré predstavujú minulé výsledky s podmienenými hustotami $r(y \mid \boldsymbol{\theta}_i)$, $i = 1, \ldots, N$, kde $\boldsymbol{\theta}_i$ sú nezávislé náhodné vektory. Pomocou týchto náhodných veličín za istých predpokladov sa dá metódami uvedenými v prameni [4] vytvoriť odhad hustoty a takéto metódy sa nazývajú ***empirické bayesovské metódy***.

3.2.1 Bayesova veta

Veta 3.1. Nech $\boldsymbol{\theta} = (\theta_1, \ldots, \theta_k)^{\mathrm{T}}$ je náhodný vektor a $q(\boldsymbol{\theta})$ je hustota tohto vektoru vzhľadom k σ-konečnej miere λ na $\Theta, \mathcal{B}(\Theta)$. Θ predstavuje neprázdnu borelovskú podmnožinu $\mathbb{R}^k$ a $\mathcal{B}(\Theta)$ označuje borelovské podmnožiny Θ. Uvažujme náhodný vektor $\mathbf{X} = (X_1, \ldots, X_n)^{\mathrm{T}}$ s podmienenou hustotou $r(\mathbf{x} \mid \boldsymbol{\theta})$ pri danom $\boldsymbol{\theta}$ vzhľadom k σ-konečnej miere ν na $(\mathbb{R}^n, \mathcal{B}^n)$. Podmienená hustota $\pi(\boldsymbol{\theta} \mid \mathbf{x})$ náhodného vektoru $\boldsymbol{\theta}$, pokiaľ je dané $\mathbf{X} = \mathbf{x}$, sa rovná

$$\pi(\boldsymbol{\theta} \mid \mathbf{x}) = \frac{r(\mathbf{x} \mid \boldsymbol{\theta}) q(\boldsymbol{\theta})}{\int_{\Theta} r(\mathbf{x} \mid \boldsymbol{\theta}) q(\boldsymbol{\theta}) d\lambda(\boldsymbol{\theta})}, \tag{3.3}$$

ak je menovateľ rôzny od nuly,

$$\pi(\boldsymbol{\theta} \mid \mathbf{x}) = 0 \tag{3.4}$$

inak.

Dôkaz: Pomocou vzťahov (3.1) a (3.2) v poznámke 3.1. prislúchajúcej definícii podmienenej hustoty máme, že

$$\pi(\boldsymbol{\theta} \mid \mathbf{x}) = \frac{f(\mathbf{x}, \boldsymbol{\theta})}{f_{\mathbf{X}}(\mathbf{x})} = \frac{r(\mathbf{x} \mid \boldsymbol{\theta}) q(\boldsymbol{\theta})}{f_{\mathbf{X}}(\mathbf{x})},$$

kde $f(\mathbf{x}, \boldsymbol{\theta})$ je združená hustota vektoru $(\mathbf{X}^{\mathrm{T}}, \boldsymbol{\theta}^{\mathrm{T}})^{\mathrm{T}}$ vzhľadom k $\nu \times \lambda$. Navyše marginálna hustota $f_{\mathbf{X}}(\mathbf{x})$ náhodného vektoru $\mathbf{X}$ sa spočíta zo združenej hustoty $f(\mathbf{x}, \boldsymbol{\theta})$ a to

$$f_{\mathbf{X}}(\mathbf{x}) = \int_{\Theta} f(\mathbf{x}, \boldsymbol{\theta}) d\lambda(\boldsymbol{\theta}) = \int_{\Theta} r(\mathbf{x} \mid \boldsymbol{\theta}) q(\boldsymbol{\theta}) d\lambda(\boldsymbol{\theta}).$$

V prípade, že je menovateľ vo vzorci (3.3) rovný nule, teda

$$\int_{\Theta} r(\mathbf{x} \mid \boldsymbol{\theta}) q(\boldsymbol{\theta}) d\lambda(\boldsymbol{\theta}) = 0,$$

tak s ohľadom na nezápornosť hustôt $r(\mathbf{x} \mid \boldsymbol{\theta})$ a $q(\boldsymbol{\theta})$ je nutne i

$$r(\mathbf{x} \mid \boldsymbol{\theta}) q(\boldsymbol{\theta}) = 0,$$

a preto platí vzťah (3.4).

$\square$

Poznámka 3.2. Obvykle sa vyjadruje $\pi(\boldsymbol{\theta} \mid \mathbf{x})$ v tvare

$$\pi(\boldsymbol{\theta} \mid \mathbf{x}) = ar(\mathbf{x} \mid \boldsymbol{\theta})q(\boldsymbol{\theta})$$

s tým, že $a = 1 / \int_{\Theta} r(\mathbf{x} \mid \boldsymbol{\theta})q(\boldsymbol{\theta})d\lambda(\boldsymbol{\theta})$ a a sa označuje ako normujúca konštanta.

3.2.2 Význam Bayesovej vety

V bayesovskej analýze má táto veta kľúčové postavenie. Je to kvôli tomu, že umožňuje získať závery o parametri $\boldsymbol{\theta}$, a to prostredníctvom ***aposteriórnej hustoty*** $\pi(\boldsymbol{\theta} \mid \mathbf{x})$, ktorá je hustotou parametra $\boldsymbol{\theta}$ po realizácii $\mathbf{X}$. Klasický prístup využíva k záverom o tomto parametri len podmienenú hustotu $r(\mathbf{x} \mid \boldsymbol{\theta})$. Bayesova veta teda hovorí, že k záverom o $\boldsymbol{\theta}$ využijeme informáciu plynúcej z realizácie $\mathbf{X}$ a takisto apriórnu informáciu parametra $\boldsymbol{\theta}$ zachytenej v apriórnej hustote $q(\boldsymbol{\theta})$, ktoré sú zahrnuté v aposteriórnej hustote.

3.3 Konjugované systémy apriórnych hustôt

Uvažujme náhodný výber $X_1, \ldots, X_n$ z rozdelenia, ktoré má vzhľadom k σ-konečnej miere ν hustotu $r(\mathbf{x} \mid \boldsymbol{\theta})$ a parameter $\boldsymbol{\theta}$ patrí do neprázdnej borelovskej množiny Θ v $\mathbb{R}^k$. Nech $q(\boldsymbol{\theta})$ je apriórna hustota parametra vzhľadom k σ-konečnej miere λ a nech nastala taká realizácia vektoru $\mathbf{X} = \mathbf{x} = (x_1, \ldots, x_n)^{\mathrm{T}}$, že platí

$$0 < \int_{\Theta} r(x_1 \mid \boldsymbol{\theta}) \ldots r(x_n \mid \boldsymbol{\theta})q(\boldsymbol{\theta})d\lambda(\boldsymbol{\theta}) < \infty. \tag{3.5}$$

Z Bayesovej vety plynie, že aposteriórna hustota je v tvare

$$\pi(\boldsymbol{\theta} \mid \mathbf{x}) = ar(x_1 \mid \boldsymbol{\theta}) \ldots r(x_n \mid \boldsymbol{\theta})q(\boldsymbol{\theta})$$

s konštantou $a > 0$, ktorá predstavuje normujúcu konštantu z poznámky 3.2.

Majme systém O apriórnych hustôt $q(\boldsymbol{\theta})$. Tento systém je konjugovaný s hustotami $r(\mathbf{x} \mid \boldsymbol{\theta})$, pokiaľ pri ľubovoľnom dostatočne veľkom rozsahu výberu n a pri ľubovoľných hodnotách náhodného výberu $X_1 = x_1, \ldots, X_n = x_n$, ktoré splňujú podmienku (3.5), aposteriórna hustota $\pi(\boldsymbol{\theta} \mid \mathbf{x})$ takisto patrí do systému O. To znamená, že systém apriórnych hustôt je konjugovaný s hustotami $r(\mathbf{x} \mid \boldsymbol{\theta})$, ak každá prislúchajúca aposteriórna hustota je rovnakého typu ako apriórna. Konjugované systémy apriórnych rozdelení teda predstavujú systémy, ktoré zahrňujú apriórne i aposteriórne hustoty.

V ďalšej časti tejto podkapitoly sa budeme zaoberať metódou konštrukcie konjugovaných systémov apriórnych hustôt a uvedieme stručný prehľad týchto systémov, ktoré prislúchajú najčastejšie používaným hustotám $r(\mathbf{x}\,|\,\boldsymbol{\theta})$. Táto metóda využíva ku konštrukcii postačujúcu štatistiku, preto je v prvom rade nutné vysvetliť pojem štatistika.

Definícia 3.2. Uvažujme náhodný vektor $\mathbf{X} = (X_1, \ldots, X_n)^{\mathrm{T}}$, ktorý má vzhľadom k σ-konečnej miere μ hustotu $f(\mathbf{x}, \boldsymbol{\theta})$ a neznámy parameter $\boldsymbol{\theta} = (\theta_1, \ldots, \theta_m)^{\mathrm{T}}$. ***Štatistika*** $\mathbf{S}$ je náhodný vektor pozostávajúci z merateľných funkcií $S_1(\mathbf{x}), \ldots, S_k(\mathbf{x})$ v tvare $\mathbf{S} = [S_1(\mathbf{X}), \ldots, S_k(\mathbf{X})]^{\mathrm{T}}$.

Poznámka 3.3. Tento vektor slúži na riešenie úloh odhadu parametrov $\boldsymbol{\theta}$, kde na základe $\mathbf{X}$ je cieľom získať čo najlepší odhad. Môže sa to týkať, či už bodového alebo intervalového odhadu, pri ktorom sa snažíme nájsť množinu, ktorá s pomerne vysokou pravdepodobnosťou pokrýva hodnotu skutočného parametra. Tieto merateľné funkcie $S_i(\mathbf{x})$ sa volia tak, aby vektor $\mathbf{S}$ získal oproti $\mathbf{X}$ nižšiu dimenziu, ale aby nestratil žiadnu informáciu o parametri $\boldsymbol{\theta}$. Ak by štatistika mala sama slúžiť ako odhad parametra, musí byť splnená rovnosť $m = k$. V ostatných prípadoch získame hľadaný odhad $\hat{\boldsymbol{\theta}}$ ďalšími krokmi práve pomocou zložiek štatistiky $\mathbf{S}$.

Definícia 3.3. Povieme, že $\mathbf{S}$ je ***postačujúca štatistika*** pre parameter $\boldsymbol{\theta}$, pokiaľ existujú nezáporné merateľné funkcie $h(\mathbf{x})$ a $g(\mathbf{s}, \boldsymbol{\theta})$ také, že skoro všade vzhľadom k μ platí

$$f(\mathbf{x}, \boldsymbol{\theta}) = g[\mathbf{S}(\mathbf{x}), \boldsymbol{\theta}]h(\mathbf{x}). \tag{3.6}$$

Ak sa dá hustota $f(\mathbf{x}, \boldsymbol{\theta})$ vyjadriť v tvare (3.6), tak sa používa výraz, že hustota sa dá faktorizovať.

Triviálnym prípadom je štatistika v tvare $\mathbf{S} = (X_1, \ldots, X_n)^{\mathrm{T}}$. Je zrejmé, že táto štatistika je postačujúca, pretože $(s_1, \ldots, s_k)^{\mathrm{T}} = (x_1, \ldots, x_n)^{\mathrm{T}}$, teda $k = n$ a stačí položiť $g(\mathbf{s}, \boldsymbol{\theta}) = f(\mathbf{s}, \boldsymbol{\theta})$ a $h(\mathbf{x}) = 1$.

3.3.1 Konštrukcia konjugovaných systémov apriórnych hustôt

Tvrdenie 3.1. Máme náhodný výber $X_1, \ldots, X_n$, ktorý pri danej hodnote $\boldsymbol{\theta}$ má združenú hustotu $r_n(\mathbf{x}\,|\,\boldsymbol{\theta})$ vzhľadom k súčinovej miere $\nu \times \ldots \times \nu$. Nech $\exists n_0 \in \mathbb{N}$ také, že $\forall n \geq n_0$ je $\mathbf{S}_n(X_1, ..., X_n)$ k-rozmernou postačujúcou štatistikou systému hustôt $r_n(\mathbf{x}\,|\,\boldsymbol{\theta})$ pre parameter $\boldsymbol{\theta}$, kde $k \in \mathbb{N}$ a nezávisí na n. Definícia 3.3. hovorí, že $\forall n \geq n_0$ existujú nezáporné merateľné funkcie $g_n(\mathbf{s},\boldsymbol{\theta})$ a $h_n(\mathbf{x})$ také, že spĺňajú vzťah (3.6) a teda

$$r_n(\mathbf{x}\,|\,\boldsymbol{\theta}) = g_n[\mathbf{S}_n(\mathbf{x}),\boldsymbol{\theta}]h_n(\mathbf{x}). \tag{3.7}$$

Vytvorme množinu

$$P_n = \{\mathbf{s} : \mathbf{s} = S_n(\mathbf{X})\}.$$

Množina P_n obsahuje všetky body $\mathbf{s} \in \mathbb{R}^k$, ktoré môže postačujúca štatistika $\mathbf{S}_n(\mathbf{X})$ nadobudnúť. Nech $\forall \mathbf{s} \in P_n$ platí, že

$$0 < \int_\Theta g_n(\mathbf{s},\boldsymbol{\theta})d\lambda(\boldsymbol{\theta}) < \infty.$$

Označme

$$c_{n,\mathbf{s}} = \frac{1}{\int_\Theta g_n(\mathbf{s},\boldsymbol{\theta})d\lambda(\boldsymbol{\theta}) < \infty}. \tag{3.8}$$

Potom systém

$$\{f_{n,\mathbf{s}}(\boldsymbol{\theta}),\ \mathbf{s} \in P_n,\ n \geq n_0\}, \tag{3.9}$$

kde

$$f_{n,\mathbf{s}}(\boldsymbol{\theta}) = g_n(\mathbf{s},\boldsymbol{\theta})c_{n,\mathbf{s}}, \tag{3.10}$$

je konjugovaný systém s $\{r_n(\mathbf{x}\,|\,\boldsymbol{\theta}),\ \boldsymbol{\theta} \in \Theta\}$.

Uvažujme teraz dva nezávislé náhodné výbery $X_1, ..., X_n$ a $Y_1, ..., Y_m$ z rozdelenia s hustotou $r(x\,|\,\boldsymbol{\theta})$. Dokážeme platnosť predchádzajúceho tvrdenia. Nech teda $n \geq n_0$ a $m \geq n_0$. Združená hustota $r_{n+m}(\mathbf{x},\mathbf{y}\,|\,\boldsymbol{\theta})$ oboch výberov je v tvare

$$r_n(\mathbf{x}\,|\,\boldsymbol{\theta})r_m(\mathbf{y}\,|\,\boldsymbol{\theta}),$$

keďže

$$r_n(\mathbf{x}\,|\,\boldsymbol{\theta}) = \prod_{i=1}^{n} r(x_i\,|\,\boldsymbol{\theta}) \text{ a } r_m(\mathbf{y}\,|\,\boldsymbol{\theta}) = \prod_{j=1}^{m} r(y_j\,|\,\boldsymbol{\theta}).$$

Preto nastáva rovnosť

$$g_{n+m}[\mathbf{S}_{n+m}(\mathbf{x},\mathbf{y}),\boldsymbol{\theta}]h_{n+m}(\mathbf{x},\mathbf{y}) = g_n[\mathbf{S}_n(\mathbf{x}),\boldsymbol{\theta}]g_m[\mathbf{S}_m(\mathbf{y}),\boldsymbol{\theta}]h_n(\mathbf{x})h_m(\mathbf{y}). \tag{3.11}$$

Kvôli jednoduchosti predpokladajme, že $\forall n \geq n_0$ je $h_n(\mathbf{x}) > 0$. Nech $\mathbf{s} \in P_m$, potom $\exists \mathbf{y} \in \mathbb{R}^m$ také, že $\mathbf{s} = S_m(\mathbf{y})$. Zvoľme za apriórnu hustotu parametra $\boldsymbol{\theta}$ funkciu $f_{m,\mathbf{s}}(\boldsymbol{\theta})$. Z Bayesovej vety vieme, že aposteriórna hustota sa spočíta vzorcom

$$\pi(\boldsymbol{\theta} \mid \mathbf{x}) = a r_n(\mathbf{x} \mid \boldsymbol{\theta}) f_{m,\mathbf{s}}(\boldsymbol{\theta}),$$

kde a je normujúca konštanta z poznámky 3.2. Využitím vzťahov (3.7) a (3.10) získavame aposteriórnu hustotu v tvare

$$\pi(\boldsymbol{\theta} \mid \mathbf{x}) = a g_n[\mathbf{S}_n(\mathbf{x}), \boldsymbol{\theta}] h_n(\mathbf{x}) g_m(\mathbf{s}, \boldsymbol{\theta}) c_{m,\mathbf{s}} =$$

$$= a c_{m,\mathbf{s}} \frac{1}{h_m(\mathbf{y})} g_n[\mathbf{S}_n(\mathbf{x}), \boldsymbol{\theta}] g_m[\mathbf{S}_m(\mathbf{y}), \boldsymbol{\theta}] h_n(\mathbf{x}) h_m(\mathbf{y}).$$

Ďalej z (3.11) dostávame

$$\pi(\boldsymbol{\theta} \mid \mathbf{x}) = b g_{n+m}(\mathbf{t}, \boldsymbol{\theta}) h_{n+m}(\mathbf{x}, \mathbf{y})$$

s vektorom $\mathbf{t} = \mathbf{S}_{n+m}(\mathbf{x}, \mathbf{y})$ a konštantou $b = a\left[c_{m,\mathbf{s}} h_m(\mathbf{y})\right]^{-1}$.

Týmto sme však zatiaľ stále nedokázali platnosť tvrdenia, preto si teraz rozpíšeme výraz $b h_{n+m}(\mathbf{x}, \mathbf{y})$ ako

$$b h_{n+m}(\mathbf{x}, \mathbf{y}) = a c_{m,\mathbf{s}} \frac{1}{h_m(\mathbf{y})} h_{n+m}(\mathbf{x}, \mathbf{y}) = a c_{m,\mathbf{s}} h(\mathbf{x}).$$

Keďže z Bayesovej vety vieme, že

$$a = \frac{1}{\int_\Theta r_n(\mathbf{x} \mid \boldsymbol{\theta}) f_{m,\mathbf{s}}(\boldsymbol{\theta}) d\lambda(\boldsymbol{\theta})},$$

tak pomocou (3.7) a (3.10) máme rovnosť

$$a c_{m,\mathbf{s}} h_n(\mathbf{x}) = \frac{1}{\int_\Theta g_{n+m}[\mathbf{S}_{n+m}(\mathbf{x}, \mathbf{y}), \boldsymbol{\theta}] d\lambda(\boldsymbol{\theta})},$$

čo vzťah (3.8) definuje ako

$$\frac{1}{\int_\Theta g_{n+m}[\mathbf{S}_{n+m}(\mathbf{x}, \mathbf{y}), \boldsymbol{\theta}] d\lambda(\boldsymbol{\theta})} = c_{n+m,\mathbf{t}}.$$

Konečne teda dostávame aposteriórnu hustotu v tvare

$$\pi(\boldsymbol{\theta} \mid \mathbf{x}) = g_{n+m}(\mathbf{t}, \boldsymbol{\theta}) c_{n+m,\mathbf{t}}, \mathbf{t} \in P_{n+m}$$

a dokázali sme, že aposteriórna hustota π patrí do systému (3.10). Tento systém sa zvykne označovať ako ***prirodzený konjugovaný systém***.

3.3.2 Prehľad konjugovaných systémov apriórnych hustôt

Touto metódou dokážeme získať konjugované systémy apriórnych hustôt. Teraz uvedieme také konjugované systémy, ktoré prináležia niektorým základným jednorozmerným systémom rozdelení $r(\mathbf{x}\,|\,\theta),\ \theta \in \Theta$, ktoré sme si uviedli v podkapitole 1.2. Zaveďme parameter s, ktorý reprezentuje $\sum_{i=1}^{n} x_i$.

Alternatívne rozdelenie

Ako prvé berme v úvahu Alternatívne rozdelenie s neznámym parametrom $\theta \in (0,1)$. Hustotou náhodného vektoru $\mathbf{X} = (X_1, ..., X_n)^{\mathrm{T}}$ vzhľadom k čítacej miere je

$$r_n(\mathbf{x}\,|\,\theta) = \theta^s (1-\theta)^{n-s}.$$

Položme

$$\mathbf{S}_n(\mathbf{x}) = \sum_{i=1}^{n} X_i,\ g_n(s,\theta) = \theta^s (1-\theta)^{n-s},\ h_n(\mathbf{x}) = 1.$$

Keďže x_i nadobúda hodnoty 1 alebo 0, tak $P_n = \{0,\ 1,\ ...,\ n\}$ a z (3.10) vyplýva, že

$$f_{n,s}(\theta) = \frac{1}{\mathrm{B}(s+1, n-s+1)} \theta^s (1-\theta)^{n-s} \text{ pre } n = 1,\ 2,\ ... \text{ a } s = 0,\ 1,\ ...,\ n.$$

V tomto prípade je prirodzený konjugovaný systém systémom beta rozdelení s parametrami $(s+1, n-s+1)$. Pokiaľ je apriórna hustota parametra θ v tvare

$$q(\theta) = \frac{1}{\mathrm{B}(\alpha,\beta)} \theta^{\alpha-1} (1-\theta)^{\beta-1},\ 0 < \theta < 1,\ \alpha > 0,\ \beta > 0,$$

potom vyjadríme aposteriórnu hustotu ako

$$\pi_n(\mathbf{x}\,|\,\theta) = \frac{1}{\mathrm{B}(\alpha+s, \beta+n-s)} \theta^{\alpha+s-1} (1-\theta)^{\beta+n-s-1}.$$

Binomické rozdelenie

Združená hustota náhodného vektoru $\mathbf{X}$ je vzhľadom k čítacej miere

$$r_n(\mathbf{x}\,|\,\theta) = \prod_{i=1}^{n} \binom{m}{x_i} \theta (1-\theta)^{mn-s}.$$

Postačujúca štatistik a funkcie $g_n,\ h_n$ sú určené ako

$$\mathbf{S}_n(\mathbf{x}) = \sum_{i=1}^{n} X_i,\ g_n(s,\theta) = \theta^s (1-\theta)^{mn-s},\ h_n(\mathbf{x}) = \prod_{i=1}^{n} \binom{m}{x_i}$$

a množina $P_n = \{0,\ 1,\ ...,\ mn\}$. Preto z (3.10) máme

$$f_{n,s}(\theta) = \frac{1}{\mathrm{B}(s+1, mn-s+1)} \theta^s (1-\theta)^{mn-s}, \; n = 1, 2, \ldots \text{ a } s = 0, 1, \ldots, mn$$

a ak apriórna hustota parametra θ je beta rozdelenie s parametrami (α, β), potom aposteriórna hustota je beta rozdelenie s parametrami $(\alpha + s, \beta + mn - s)$ v tvare

$$\pi_n(\mathbf{x} \mid \theta) = \frac{1}{\mathrm{B}(\alpha + s, \beta + mn - s)} \theta^{\alpha+s-1} (1-\theta)^{\beta+mn-s-1}.$$

Poissonovo rozdelenie

Prirodzeným konjugovaným systémom v Poissonovom rozdelení s parametrom θ je $\Gamma(m, s)$, $m = 1, 2, \ldots$ a $s = 1, 2, \ldots$, kde postačujúcou štatistikou je

$$\mathbf{S}_n(\mathbf{x}) = \sum_{i=1}^{n} X_i,$$

a preto $P_n = \{0, 1, 2, \ldots\}$. V praxi sa vyžaduje len to, aby tieto parametre spĺňali nerovnosti $m > 0$, $s > 0$. Pre apriórne rozdelenie $\Gamma(a, p)$ je aposteriórne rozdelenie $\Gamma(a + n, p + s)$.

Rovnomerné rozdelenie

Uvažujme rovnomerné rozdelenie s parametrami θ_1 a θ_2. Ako postačujúce štatistiky položme

$$\min_{1 \le i \le n} X_i, \; \max_{1 \le i \le n} X_i,$$

a preto $P_n = (s_1, s_2), s_1 < s_2$, kde $s_1 = \min_{1 \le i \le n} x_i$ a $s_2 = \max_{1 \le i \le n} x_i$. Systém dvojrozmerných Paretových rozdelení s parametrami (s_1, s_2, m), $s_1 < s_2$, $m = 1, 2, \ldots$ je prirodzeným konjugovaným systémom. Aposteriórne rozdelenie je Paretovo rozdelenie s parametrami $\min(r_1, X_1, \ldots., X_n), \max(r_2, X_1, \ldots., X_n), m+n$, pokiaľ je apriórne rozdelenie parametra θ rovnakého typu s parametrami (r_1, r_2, m).

Exponencionálne rozdelenie

V exponencionálnom rozdelení s parametrom $\theta \in (0, \infty)$ položme rovnako

$$\mathbf{S}_n(\mathbf{x}) = \sum_{i=1}^{n} X_i.$$

Keďže ide o spojité rozdelenie, kde X_i nadobúda kladné hodnoty, dostávame, že $P_n = (0, \infty)$. Prirodzeným konjugovaným systémom je systém gama rozdelení s parametrami (s, m) pre $s > 0$, $m = 1, 2, \ldots$ Ak teda zvolíme za apriórne rozdelenie $\Gamma(a, p)$, je aposteriórnym rozdelením parametra θ rozdelenie $\Gamma(a + s, p + n)$.

Normálne rozdelenie

Posledným systémom rozdelenia, ktorého prirodzený konjugovaný systém hľadáme, je normálne rozdelenie so známym rozptylom σ_0^2. Množina $P_n = \mathbb{R}$, pretože postačujúca štatistika je daná ako $\sum_{i=1}^{n} X_i$. Systém normálnych rozdelení s parametrami $\eta_0 \in \mathbb{R}$ a $\frac{\sigma_0^2}{n} > 0$ pre $n = 1, 2, ...,$ je príslušný prirodzený konjugovaný systém. Avšak v praxi sa namiesto neho využíva väčší systém, ktorý tvorí množina hustôt všetkých normálnych rozdelení s parametrami $\eta_0 = \eta_1 \in \mathbb{R}$ a $\sigma_1^2 > 0.$ V prípade, že za apriórne rozdelenie zvolíme $\mathrm{N}(\eta_1, \sigma_1^2)$, získame aposteriórne rozdelenie $\mathrm{N}(\eta_2, \sigma_2^2)$, kde príslušné parametre sú v tvare

$$\eta_2 = \frac{\eta_0 \sigma_0^2 + s\sigma_1^2}{\sigma_0^2 + n\sigma_1^2}, \ \sigma_2^2 = \frac{\sigma_0^2 \sigma_1^2}{\sigma_0^2 + n\sigma_1^2}.$$

3.4 Princíp neurčitosti

Doteraz sme sa zaoberali prípadmi, kde máme dostatok informácií o parametri $\boldsymbol{\theta}$ a poznáme apriórne rozdelenie parametra $\boldsymbol{\theta}$. ***Princíp neurčitosti*** sa uplatňuje práve vtedy, keď apriórne rozdelenie nepoznáme a napriek tomu chceme získať aposteriórne rozdelenie.

Nech opäť $\mathbf{X}$ predstavuje náhodný vektor s podmienenou hustotou $r(\mathbf{x} \mid \boldsymbol{\theta})$ pri danej hodnote parametra $\boldsymbol{\theta} \in \Theta$, kde $\Theta \in \mathcal{B}^k$. O náhodnom vektore $\boldsymbol{\theta}$, okrem toho, že patrí do množiny Θ, nemáme žiadne informácie, preto nastáva problém ako zvoliť apriórne rozdelenie. V takomto prípade podľa princípu neurčitosti volíme za apriórne rozdelenie $\boldsymbol{\theta}$ rovnomerné rozdelenie na množine Θ. Z tohto dôvodu sa apriórna hustota volí ako nejaká kladná konštanta. Keďže hustota $cq_0(\boldsymbol{\theta})$, kde $c > 0$ a $q_0(\boldsymbol{\theta})$ je označenie apriórnej hustoty, zahrnutím konštanty c do normujúcej konštanty z poznámky 3.2., nijako neovplyvní aposteriórnu hustotu vo vzorcoch (3.3) a (3.4). Môžeme bez ohľadu na veľkosť množiny Θ položiť $q_0(\boldsymbol{\theta})$ hodnote 1 na Θ. Pokiaľ je množina Θ nanajvýš spočetná, chápeme q_0 ako hustotu vzhľadom k čítacej miere. V prípade, že je Lebesgueova miera Θ kladná, hustotu q_0

považujeme za hustotu vzhľadom k Lebesgueovej miere. Ak nastane prípad, že Θ je nekonečne spočetná alebo ak je Lebesgueova miera Θ nekonečná, tak povieme, že hustota q_0 je nevlastná. Rovnako však i v tomto prípade položíme $q_0(\boldsymbol{\theta}) = 1$.

Uvažujme náhodný výber $X_1, \ldots, X_n$ z alternatívneho rozdelenia s parametrom θ patriaceho do intervalu $(0,1)$. Predpokladajme, že o θ nemáme žiadne informácie. Potom podľa princípu neurčitosti zvoľme za apriórnu hustou $q_0(\theta)$ vzhľadom k Lebesgueovej miere

$$q_0(\theta) = 1, \ \ \theta \in (0,1).$$

Ak takto zvolíme apriórnu hustotu, získame aposteriórne rozdelenie v podobe beta rozdelenia s parametrami $(\sum_{i=1}^{n} x_i + 1, n - \sum_{i=1}^{n} x_i + 1)$ a z tohto dôvodu je aposteriórna hustota $\pi_0(\theta \mid \mathbf{x})$ v tvare

$$\pi_0(\theta \mid \mathbf{x}) = \frac{1}{\mathrm{B}(\sum_{i=1}^{n} x_i + 1, n - \sum_{i=1}^{n} x_i + 1)} \theta^{\sum_{i=1}^{n} x_i} (1-\theta)^{n - \sum_{i=1}^{n} x_i}.$$

Na jednej strane je využitie princípu neurčitosti v prípade nedostatku informácií o parametri $\boldsymbol{\theta}$ výhodné, avšak na strane druhej jeho použitie v niektorých prípadoch nie je veľmi vhodné. Existujú totižto prípady, kedy použitie princípu neurčitosti vedie k sporným záverom.

Majme náhodný výber $X_1, \ldots, X_n$ z normálneho rozdelenia, kde σ^2 je známy kladný parameter a η je parameter neznámy patriaci do množiny reálnych čísel $\mathbb{R}$. Pre jednoduchosť označme

$$\frac{1}{n}\sum_{i=1}^{n} X_i = \bar{X}.$$

Pokiaľ pri zvolení apriórnej hustoty parametra η sa budeme riadiť princípom neurčitosti, tak získame aposteriórnu hustotu, ktorá prináleží $\mathrm{N}(\bar{X}, n^{-1})$. Z teórie pravdepodobnosti plynie, že náhodná veličina Y, ktorá vznikla ako

$$Y = \bar{X} - \eta,$$

má podmienené rozdelenie pri danom $\bar{X}$ normálne s parametrami $(0, \frac{\sigma^2}{n})$ rovnako tak aj pri danom η a obe dvojice náhodných veličín $Y, \bar{X}$ a Y, η sú nezávislé. Vzhľadom na vzťahy (1.2) a (1.4) vieme, že v takomto prípade platí

$$\operatorname{var}\bar{X} = \operatorname{var}\eta + \operatorname{var}Y, \quad \operatorname{var}\eta = \operatorname{var}\bar{X} + \operatorname{var}Y$$

a $\operatorname{var}Y = 0$ je dôsledok tejto sústavy rovníc. To, že rozptyl náhodnej veličiny Y je nulový, je spor s tým, že táto náhodná veličina má rozdelenie $\mathrm{N}(0, \frac{\sigma^2}{n})$ pre $\sigma^2 > 0$. Označme si tento spor ako $(*)$ [1].

Nech $X_1, ..., X_n$ je náhodný výber z Poissonového rozdelenia s kladným parametrom θ. Zvoľme apriórnu hustotu $q_0(\theta)$ podľa princípu neurčitosti. Potom aposteriórna hustota je hustotou gama rozdelenia s parametrami $(n, \sum_{i=1}^{n} x_i + 1)$. Zmeňme teraz parametrizáciu modelu, to znamená, že namiesto parametra θ vezmime parameter $\eta = \theta^{1/2}$. Keďže η a θ vnímame ako náhodné veličiny, tak využitím tvrdenia 1.6., s ohľadom na požiadavku kladnosti η v Poissonovom rozdelení, získavame nekonštantnú apriórnu hustotu nového parametra v tvare

$$q_0(\eta) = \eta, \quad \eta > 0.$$

Tento tvar hustoty vedie k paradoxnému záveru, pretože o hodnotách parametra η máme určitú informáciu napriek tomu, že o hodnotách θ nevieme nič. Príslušná aposteriórna hustota je

$$\pi_0(\eta \mid \mathbf{x}) = \frac{2n^{\sum_{i=1}^{n} x_i + 1}}{\Gamma(\sum_{i=1}^{n} x_i + 1)} e^{-\eta^2 n} \eta^{2\sum_{i=1}^{n} x_i + 1} \quad \text{pre } \eta > 0.$$

Pokiaľ teraz zvolíme pre η rovnomernú apriórnu hustotu podľa princípu neurčitosti, teda

$$q_0^*(\eta) = 1, \quad \eta > 0,$$

tak potom príslušnou aposteriórnou hustotou je

$$\pi_0^*(\eta \mid \mathbf{x}) = \frac{2n^{\sum_{i=}^{n} x_i + 1/2}}{\Gamma(\sum_{i=1}^{n} x_i + 1/2)} e^{-\eta^2 n} \eta^{2\sum_{i=1}^{n} x_i} \quad \text{pre } \eta > 0.$$

Získali sme inú aposteriórnu hustotu ako pri použití apriórnej hustoty, pri ktorej máme o parametri η určitú informáciu. Keďže oba postupy výpočtu aposteriórnych hustôt π_0, π_0^* sú z logického hľadiska rovnocenné a tieto hustoty sú navzájom rôzne, vzniká paradoxný záver. Práve táto skutočnosť vedie k tomu, že by bolo lepšie

[1] K sporu $(*)$ sa vrátime na strane 28.

neriadiť sa princípom neurčitosti, ale za apriórne rozdelenie voliť také, ktoré nezáleží na počiatočnej parametrizácií modelu. O tom ako vyriešiť tento problém v prípade, že λ je Lebesgueova miera a $\lambda(\Theta) > 0$, už vysvetľuje nasledujúca podkapitola.

3.5 Jeffreysova apriórna hustota

Nech náhodný vektor $\mathbf{X} = (X_1, \ldots, X_n)^{\mathrm{T}}$ má vzhľadom k σ -konečnej miere μ hustotu $r(\mathbf{x}\,|\,\boldsymbol{\theta})$ s parametrom $\boldsymbol{\theta} = (\theta_1, \ldots, \theta_m)^{\mathrm{T}}$. Predtým, než si ***Jeffreysovu apriórnu husotu*** predstavíme, definujme pojmy, ktoré sú potrebné k jej zavedeniu.

Definícia 3.4. O štvorcovej matici $\mathbf{A}$, pre ktorú platí

$$\mathbf{u} \neq 0 \Rightarrow \mathbf{u}^{\mathrm{T}}\mathbf{A}\mathbf{u} > 0, \tag{3.12}$$

hovoríme, že je pozitívne definitná.

Poznámka 3.4. Každú symetrickú pozitívne definitnú maticu $\mathbf{A}$ je možné vyjadriť v tvare

$$\mathbf{A} = \mathbf{B}^{\mathrm{T}}\mathbf{B},$$

kde $\mathbf{B}$ je dolná trojuholníková matica.

Definícia 3.5. Systém hustôt $\{r(\mathbf{x}\,|\,\boldsymbol{\theta})\},\ \boldsymbol{\theta} \in \Theta\}$ nazveme ***regulárny***, pokiaľ je splnené:

- množina Θ je neprázdna a otvorená v $\mathbb{R}^k$,
- množina $A = \{\mathbf{x} : r(\mathbf{x}\,|\,\boldsymbol{\theta}) > 0\}$ nie je závislá na $\boldsymbol{\theta}$,
- pre skoro všetky $\mathbf{x} \in A$ vzhľadom k μ existuje konečná parciálna derivácia $\dfrac{\partial r(\mathbf{x}\,|\,\boldsymbol{\theta})}{\partial \theta_i}$, $i = 1, 2, \ldots, m$,
- pre všetky i a $\boldsymbol{\theta} \in \Theta$ platí $\displaystyle\int_A \frac{\partial r(\mathbf{x}\,|\,\boldsymbol{\theta})}{\partial \theta_i} d\mu(\boldsymbol{\theta}) = 0$,
- $\displaystyle \mathbf{J}_{ij}(\boldsymbol{\theta}) = \int_A \frac{\partial r(\mathbf{x}\,|\,\boldsymbol{\theta})}{\partial \theta_i} \frac{\partial r(\mathbf{x}\,|\,\boldsymbol{\theta})}{\partial \theta_j} \frac{1}{r(\mathbf{x}\,|\,\boldsymbol{\theta})^2} r(\mathbf{x}\,|\,\boldsymbol{\theta}) d\mu(\boldsymbol{\theta})$ existuje a je konečný pre $\forall i, j$,
- matica $\mathbf{J}(\boldsymbol{\theta}) = J_{ij}(\boldsymbol{\theta})_{\,i,j=1,\ldots m}$ je pre $\forall \boldsymbol{\theta} \in \Theta$ pozitívne definitná. Táto matica sa zvykne označovať ako ***Fisherova informačná matica***.

3.5.1 Jeffreysova veta

Veta 3.2. Majme hustotu $r(\mathbf{x}|\boldsymbol{\theta})$ vzhľadom k σ -konečnej miere μ náhodného vektoru $\mathbf{X}=(X_1, ..., X_n)^{\mathrm{T}}$. Nech systém $\{r(\mathbf{x}|\boldsymbol{\theta})\}, \boldsymbol{\theta}\in\Theta\}$ je regulárny a má Fisherovu informačnú maticu $\mathbf{J}(\boldsymbol{\theta})$. Nech je splnené

$$0<\int_{\Theta} r(\mathbf{x}|\boldsymbol{\theta})\,|\,\mathbf{J}(\boldsymbol{\theta})\,|^{1/2}\,d\boldsymbol{\theta}<\infty$$

a položme

$$c=(\int_{\Theta} r(\mathbf{x}|\boldsymbol{\theta})\,|\,\mathbf{J}(\boldsymbol{\theta})\,|^{1/2}\,d\boldsymbol{\theta})^{-1}.$$

Nech $\mathbf{H}$ je regulárne prosté zobrazenie množiny Θ na $\Theta^*\in\mathcal{B}^k$. Označme $\boldsymbol{\theta}^*=\mathbf{H}(\boldsymbol{\theta})$ a $r^*(\mathbf{x}|\boldsymbol{\theta}^*)=r(\mathbf{x}|\mathbf{H}^{-1}(\boldsymbol{\theta}^*))$. Získame tak regulárny systém hustôt $\{r^*(\mathbf{x}|\boldsymbol{\theta}^*)\}, \boldsymbol{\theta}^*\in\Theta^*\}$. Pokiaľ označíme $\mathbf{J}^*(\boldsymbol{\theta}^*)$ ako Fisherovu informačnú maticu tohto systému, potom pre ľubovoľnú množinu $B\subset\Theta$, ktorá patrí do $\mathcal{B}^k$, platí

$$\int_B c r(\mathbf{x}|\boldsymbol{\theta})\,|\,\mathbf{J}(\boldsymbol{\theta})\,|^{1/2}\,d\boldsymbol{\theta}=\int_{\mathbf{H}(B)} c r^*(\mathbf{x}|\boldsymbol{\theta}^*)\,|\,\mathbf{J}(\boldsymbol{\theta}^*)\,|^{1/2}\,d\boldsymbol{\theta}^*.$$

Dôkaz: [4] na strane 30 a 31.

Z vety 3.1. za predpokladov vety 3.2. vyplýva, že ak apriórnou hustotou parametra $\boldsymbol{\theta}$ je akýkoľvek kladný násobok funkcie $|\,\mathbf{J}(\boldsymbol{\theta})\,|^{1/2}$, tak potom aposteriórna hustota π má tvar $c r(\mathbf{x}|\boldsymbol{\theta})\,|\,\mathbf{J}(\boldsymbol{\theta})\,|^{1/2}$ a je pravdepodobnostnou hustotou. Apriórnu hustotu $k\,|\,\mathbf{J}(\boldsymbol{\theta})\,|^{1/2}$, kde $k>0$, označujeme ako Jeffreysovu apriórnu hustotu. Veta 3.2. hovorí, že pri Jeffreysovej voľbe apriórnej hustoty parametrov $\boldsymbol{\theta}$ a $\boldsymbol{\theta}^*$ sú obe aposteriórne pravdepodobnosti rovnaké, preto nenastávajú paradoxné závery ako pri využití princípu neurčitosti.

Pre príklad si teraz vezmime Poissonnovo rozdelenie s parametrom λ. Jeffreysova apriórna hustota tohto parametra je

$$q(\lambda)=\lambda^{-1/2} \text{ pre } \lambda>0,\; q(\lambda)=0 \text{ inak.}$$

Hustota je nevlastná, čo pri tomto type apriórnej hustoty môže nastať. Aposteriórna hustota je $\Gamma(n,\sum_{i=1}^{n} x_i+1/2)$.

V podkapitole 3.4. sme ukázali, prečo nie je vhodné použitie princípu neurčitosti pri normálnom rozdelení, kde $\sigma^2>0$ je známe a η neznáme. Jeffreysova apriórna hustota je v tomto prípade konštantná, čo vedie takisto k sporu $(*)$.

Na záver si ešte uveďme prípad normálneho rozdelenia s neznámym η i σ^2. Jeffreysova apriórna hustota je

$$q(\eta, \sigma^{-2}) = \sigma, \;\; \eta \in \mathbb{R}, \;\; \sigma^{-2} > 0.$$

Táto hustota je nevlastná a čo sa týka aposteriórnej hustoty, tak podmienené rozdelenie η pri danom σ^{-2} je $\mathrm{N}(\overline{x}, \frac{\sigma^2}{n})$ a marginálnym rozdelením veličiny σ^{-2} je

$$\Gamma(\frac{1}{2}\sum_{i=1}^{n}(x_i - \overline{x})^2, \frac{n}{2}).$$

4 Odhady parametrov bayesovským prístupom

Táto kapitola spracováva problematiku odhadov parametrov z bayesovského pohľadu, pričom nástrojom pre ich odhadovanie sú matematické vety a definície, ktoré nájdeme v prameni [2]. V prvej podkapitole sa venujeme odhadom v modeli AR, druhá podkapitola rozoberá odhady modelu MA ako aj zložitejšieho modelu ARMA.

4.1 Odhady parametrov AR modelu

Odhady autoregresných parametrov a testovanie hypotéz o týchto parametroch sú významnou súčasťou analýzy časových radov. Aj keď sa tieto odhady dajú získať aj inými spôsobmi, ako je napríklad ***metóda najmenších štvorcov*** (viď. prameň [2]) alebo ***metóda momentov*** (viď. prameň [7]), my v tejto práci využijeme bayesovskú analýzu. Výhodou tejto voľby je, že odhady parametrov sa dajú získať relatívne ľahko a nepochybne aj to, že bayesovským prístupom získame charakteristiky s pravdepodobnostnými rozdeleniami $(\chi^2, \mathrm{t}, \mathrm{F})$. Pomocou nich dokážeme testovať významnosť parametrov. Viac o tejto problematike sa dozvieme v kapitole 5. Predtým, než sa začneme venovať samotným odhadom parametrov, je ešte nutné dodať, že slabinou tohto prístupu je malé zdôvodnenie použitej apriórnej hustoty.

Uvažujme autoregresný model s fixovaným počiatkom, čo znamená, že náhodné veličiny X_i nadobúdajú dané reálne hodnoty x_i pre $i = 1, 2, ..., n$, a že náhodné veličiny $X_{n+1}, ..., X_N$ sú vytvárané vzorcom (2.3). Upravme si teraz tento vzorec tak, aby sme získali osamotenú náhodnú veličinu X_i na ľavej strane, teda

$$X_i = b_1 X_{i-1} + ... + b_n X_{i-n} + a^{-1} Y_i \text{ pre } i = n+1, ..., N,$$

kde

$$a = d_0, \; b_j = -\frac{d_j}{d_0} \text{ pre } j = 1, ..., n.$$

Označme $\mathbf{b} = (b_1, b_2, ..., b_n)^{\mathrm{T}}$. Nasledujúca veta nám ukáže, ako nájdeme podmienenú hustotu veličín $X_{n+1}, ..., X_N$.

Veta 4.1. Pokiaľ náhodné veličiny $Y_{n+1}, \ldots, Y_N$ majú združené normované normálne rozdelenie, potom podmienená hustota náhodných veličín $X_{n+1}, \ldots, X_N$ pri daných hodnotách $x_{n+1}, \ldots, x_N$, a, $\mathbf{b}$ je rovná

$$r(x_{n+1},\ldots,x_N \mid x_1,\ldots,x_n,a,\mathbf{b}) = (2\pi)^{\frac{-(N-n)}{2}} a^{N-n} e^{\frac{-a^2}{2}\left[\sum_{j=n+1}^{N}\left(x_j-\sum_{i=1}^{n} b_i x_{j-i}\right)^2\right]}. \tag{4.1}$$

Dôkaz: Združená hustota normovaného normálneho rozdelenia náhodných veličín $Y_{n+1}, \ldots, Y_N$ je

$$(2\pi)^{\frac{-(N-n)}{2}} e^{-\frac{1}{2}\left(\sum_{j=n+1}^{N} y_j^2\right)}.$$

V tomto bode využijeme tvrdenie 1.6. z kapitoly 1, ktoré spadá do teórie transformácie náhodných vektorov. V našom prípade je

$$g_i^{-1}(\mathbf{X}) = a(X_i - b_1 X_{i-1} - \ldots - b_n X_{i-n})$$

a platí

$$\det\frac{\partial g^{-1}(\mathbf{x})}{\partial(\mathbf{x})} = \det\begin{vmatrix} \frac{\partial g_{n+1}^{-1}(\mathbf{x})}{\partial x_{n+1}} & \cdots & \frac{\partial g_{n+1}^{-1}(\mathbf{x})}{\partial x_N} \\ \vdots & \ddots & \vdots \\ \frac{\partial g_N^{-1}(\mathbf{x})}{\partial x_{n+1}} & \cdots & \frac{\partial g_N^{-1}(\mathbf{x})}{\partial x_N} \end{vmatrix}.$$

Preto

$$\det\frac{\partial g^{-1}(\mathbf{x})}{\partial(\mathbf{x})} = \det\begin{vmatrix} a & 0 & 0 & \cdots & \cdots & \cdots & 0 & 0 \\ -ab_1 & a & 0 & 0 & \cdots & \cdots & 0 & 0 \\ -ab_2 & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \vdots \\ -ab_n & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \vdots \\ 0 & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & -ab_n & \cdots & \cdots & -ab_1 & a & 0 \\ 0 & \cdots & 0 & -ab_n & \cdots & -ab_2 & -ab_1 & a \end{vmatrix}.$$

Keďže ide o o hornú trojuholníkovú maticu, ktorá má na pozíciách (i,i) parameter a pre $i = 1, \ldots, N$, nastáva rovnosť

$$\det\frac{\partial g^{-1}(\mathbf{x})}{\partial(\mathbf{x})} = a^{N-n}.$$

Následne po vyjadrení y_i v tvare vzorca (2.3) a dosadení pomocou parametrov a, $\mathbf{b}$ sme vetu dokázali.

□

Uvažujme parameter $b_0 = -1$ a položme m_{ij} do rovnosti

$$m_{ij} = \sum_{k=n+1}^{N} x_{k-i} x_{k-j},\ i, j = 0, 1, ..., n. \tag{4.2}$$

Keďže

$$\sum_{k=n+1}^{N} (x_k - \sum_{i=1}^{n} b_i x_{k-i})^2 = \sum_{k=n+1}^{N} (\sum_{i=0}^{n} b_i x_{k-i})^2 =$$

$$= \sum_{i=0}^{n} \sum_{j=0}^{n} b_i b_j \sum_{k=n+1}^{N} x_{k-1} x_{k-j} = \sum_{i=0}^{n} \sum_{j=0}^{n} b_i b_j m_{ij},$$

tak potom môžeme podmienenú hustotu (4.1) náhodných veličín X_{n+1}, ..., X_N upraviť na tvar

$$r(x_{n+1},...,x_N \mid x_1,...,x_n,a,\mathbf{b}) = (2\pi)^{\frac{-(N-n)}{2}} a^{N-n} e^{\frac{-a^2}{2}\left[\sum_{i=0}^{n}\sum_{j=0}^{n} b_i b_j m_{ij}\right]}. \tag{4.3}$$

Z dôvodu toho, že sa riadime bayesovským prístupom, budeme predpokladať, že parametre a, $\mathbf{b}$ sú náhodné veličiny. Navyše, keďže o $\mathbf{b}$ nemáme žiadne informácie, využíva sa princíp neurčitosti, preto apriórna hustota tohto vektoru v $\mathbb{R}^n$ je rovná hodnote 1. O parametri a sa predpokladá kladnosť a zároveň, že hustota náhodnej veličiny $\ln(a)$ je 1 v $\mathbb{R}$. Využitím tvrdenia 1.6. je apriórna hustota parametra a rovná a^{-1} pre $a > 0$ a z predpokladu kladnosti je táto hustota rovná 0 pre $a \leq 0$. Takisto sa predpokladá nezávislosť parametrov a, b_i pre $i = 1, ..., n$. Z tvrdenia 1.3. vyplýva, že ich združená hustota je vyjadrená v tvare a^{-1} v prípade, že $a > 0$ a 0 inak.

Veta 4.2. Vytvorme matice $\mathbf{M}$ a $\mathbf{M}^*$ nasledujúcim spôsobom. Nech

$$\mathbf{M} = (m_{ij})_{i,j=1}^{n},\ \mathbf{M}^* = (m_{ij})_{i,j=0}^{n}$$

a urobme predpoklad, že matica $\mathbf{M}^*$ splňuje definíciu 3.4. a $N > n+1$. Ak je pre $a > 0$ apriórna hustota náhodných veličín a, b_1, ..., b_n rovná výrazu a^{-1} nezávisle na x_1, ..., x_n, potom má ich aposteriórna hustota tvar

$$\pi(a,\mathbf{b} \mid \mathbf{x}) = c a^{N-n-1} e^{\frac{-a^2}{2}\left[(\mathbf{b}-\bar{\mathbf{b}})^{\mathrm{T}}\mathbf{M}(\mathbf{b}-\bar{\mathbf{b}})+m_{00}-\bar{\mathbf{b}}^{\mathrm{T}}\mathbf{M}\bar{\mathbf{b}}\right]}, \tag{4.4}$$

kde

$$\bar{\mathbf{b}} = (\bar{b}_1, ..., \bar{b}_n)^{\mathrm{T}} = \mathbf{M}^{-1}(m_{01}, ..., m_{0n})^{\mathrm{T}} \tag{4.5}$$

a $c_1 > 0$ nezávisí na a a ani $\mathbf{b}$. Modus tohto aposteriórneho rozdelenia je $(\bar{a}, \bar{\mathbf{b}})$. Hodnotu $\bar{a}$ získame zo vzťahu

$$\bar{a}^{-2} = (N-n-1)^{-1}(m_{00} - \bar{\mathbf{b}}^{\mathrm{T}}\mathbf{M}\bar{\mathbf{b}}). \tag{4.6}$$

Marginálnou aposteriórnou hustotou náhodného vektoru $\mathbf{b}$ je

$$\pi_1(\mathbf{b} \mid \mathbf{x}) = c_1\left[1 + \frac{(\mathbf{b}-\bar{\mathbf{b}})^{\mathrm{T}}\mathbf{M}(\mathbf{b}-\bar{\mathbf{b}})}{m_{00} - \bar{\mathbf{b}}^{\mathrm{T}}\mathbf{M}\bar{\mathbf{b}}}\right]^{-(N-n)/2}, \tag{4.7}$$

kde o konštante c_1 môžeme povedať, že je kladná a nezávisí na a a ani na $\mathbf{b}$.

Dôkaz: Podľa Bayesovej vety vypočítame aposteriórnu hustotu zo vzťahu (3.3). V našom prípade je menovateľ Bayesovho vzorca rovný integrálu

$$\int_{\mathbb{R}^n}\left(\int_0^\infty (2\pi)^{\frac{-(N-n)}{2}} a^{N-n} e^{\frac{-a^2}{2}\left[\sum_{i=0}^n \sum_{j=0}^n b_i b_j m_{ij}\right]} a^{-1} da\right) d\mathbf{b},$$

ktorého prevrátenú hodnotu vynásobíme členom $(2\pi)^{\frac{-(N-n)}{2}}$ a získame tak konštantu c. To, že c nezávisí na a a ani $\mathbf{b}$, plynie z toho, že po zintegrovaní sa v c nevyskytuje žiaden z parametrov a, $\mathbf{b}$. Čitateľ predstavuje súčin podmienenej hustoty (4.3) a apriórnej hustoty vektoru $(a, \mathbf{b})$, ale keďže sme už člen $(2\pi)^{\frac{-(N-n)}{2}}$ zahrnuli do konštanty c, má tvar

$$a^{N-n} e^{\frac{-a^2}{2}\left[\sum_{i=0}^n \sum_{j=0}^n b_i b_j m_{ij}\right]} a^{-1}.$$

To vysvetľuje prítomnosť člena a^{N-n-1} v aposteriórnej hustote (4.4). Položme

$$\mathbf{m}_0 = (m_{01}, \ \ldots, \ m_{0n})^{\mathrm{T}}.$$

Zo vzťahu (4.2) a spôsobu tvorby matice $\mathbf{M}^*$ vyplýva, že táto matica je symetrická a preto $m_{0k} = m_{k0}$. Potom postupnými úpravami získavame

$$\sum_{i=0}^n \sum_{j=0}^n b_i b_j m_{ij} = \sum_{i=1}^n \sum_{j=1}^n b_i b_j m_{ij} - \sum_{j=1}^n b_j m_{0j} - \sum_{i=1}^n b_i m_{i0} + d_{00} =$$
$$= \mathbf{b}^{\mathrm{T}}\mathbf{M}\mathbf{b} - \mathbf{b}^{\mathrm{T}}\mathbf{m}_0 - \mathbf{m}_0{}^{\mathrm{T}}\mathbf{b} + m_{00}$$

a po dosadení vzťahu (4.5) za $\bar{\mathbf{b}}$ dostaneme rovnosť

$$\mathbf{b}^{\mathrm{T}}\mathbf{M}\mathbf{b} - \mathbf{b}^{\mathrm{T}}\mathbf{m}_o - \mathbf{m}_0{}^{\mathrm{T}}\mathbf{b} + m_{00} = (\mathbf{b}-\bar{\mathbf{b}})^{\mathrm{T}}\mathbf{M}(\mathbf{b}-\bar{\mathbf{b}}) + m_{00} - \bar{\mathbf{b}}^{\mathrm{T}}\mathbf{M}\bar{\mathbf{b}}$$

a tým sme spolu s predošlými tvrdeniami vzorec (4.4) dokázali.

Označme vektor

$$\mathbf{b}^* = (b_0, \ \bar{b}_1, \ \ldots, \ \bar{b}_n)^{\mathrm{T}},$$

ktorý je určite vďaka $b_0 = -1$ nenulový, a keďže je z predpokladu vety matica $\mathbf{M}^*$ pozitívne definitná, tak zo vzťahu (3.12) v definícii 3.4. vyplýva, že

$$0<(\mathbf{b}^*)^{\mathrm{T}}\mathbf{M}^*\mathbf{b}^*=\sum_{i=0}^{n}\sum_{j=0}^{n}b_i^*b_j^*m_{ij}=\sum_{i=1}^{n}\sum_{j=1}^{n}\bar{b}_i\bar{b}_jm_{ij}-\sum_{j=1}^{n}\bar{b}_jm_{0j}-$$
$$-\sum_{i=1}^{n}\bar{b}_im_{i0}+m_{00}=\bar{\mathbf{b}}^{\mathrm{T}}\mathbf{M}\bar{\mathbf{b}}-\bar{\mathbf{b}}^{\mathrm{T}}\mathbf{m}_0-\mathbf{m}_0^{\mathrm{T}}\bar{\mathbf{b}}+m_{00}=m_{00}-\bar{\mathbf{b}}^{\mathrm{T}}\mathbf{M}\bar{\mathbf{b}}. \tag{4.8}$$

Z predpokladu $\mathbf{M}^*>0$ plynie i $\mathbf{M}>0.$ Zvoľme si teraz nejakú pevnú kladnú hodnotu za $a,$ potom z tvaru aposteriórnej hustoty (4.4) je vidieť, že táto hustota nadobúda maximum v $\mathbf{b}=\bar{\mathbf{b}}$ a vďaka (4.8) je výraz s exponencionálou v aposteriórnej hustote menší ako 1. Dosaďme teda za $\mathbf{b}$ vektor $\bar{\mathbf{b}}$ a získavame tak aposteriórnu hustotu v tvare

$$\pi(a,\mathbf{b}\,|\,\mathbf{x})=ca^{N-n-1}e^{\frac{-a^2}{2}\left[m_{00}-\bar{\mathbf{b}}^{\mathrm{T}}\mathbf{M}\bar{\mathbf{b}}\right]}.$$

Aposteriórnu hustotu zderivujeme podľa premennej a a položíme do rovnosti s 0, aby sme zistili pre aké a nadobúda svoje maximum. Získame tak rovnosť

$$0=c(N-n-1)a^{N-n-2}e^{\frac{-a^2}{2}\left[m_{00}-\bar{\mathbf{b}}^{\mathrm{T}}\mathbf{M}\bar{\mathbf{b}}\right]}-c(m_{00}-\bar{\mathbf{b}}^{\mathrm{T}}\mathbf{M}\bar{\mathbf{b}})a^{N-n}e^{\frac{-a^2}{2}\left[m_{00}-\bar{\mathbf{b}}^{\mathrm{T}}\mathbf{M}\bar{\mathbf{b}}\right]},$$

z ktorej po vykrátení nenulových členov nastane rovnosť

$$(N-n-1)^{-1}(m_{00}-\bar{\mathbf{b}}^{\mathrm{T}}\mathbf{M}\bar{\mathbf{b}})=a^{-2}.$$

Preznačením a na $\bar{a}$ sme dokázali (4.6). Predpoklad vety $N>n+1$ a vzťah (4.8) sú potrebné jednak pri úpravách, ktoré vedú k nájdeniu maxima hustoty, aby nedošlo ku deleniu členom s nulovou hodnotou, a rovnako tak aj z dôvodu kladnosti parametra $a,$ ktorá sa predpokladá. Čo sa týka vzorca (4.7), tak ten získame integráciou aposteriórnej hustoty (4.4) podľa parametra $a.$ A tak

$$\pi_1(\mathbf{b}\,|\,\mathbf{x})=\int_0^\infty\pi(a,\mathbf{b}\,|\,\mathbf{x})da=\int_0^\infty ca^{N-n-1}e^{\frac{-a^2}{2}\left[(\mathbf{b}-\bar{\mathbf{b}})^{\mathrm{T}}\mathbf{M}(\mathbf{b}-\bar{\mathbf{b}})+m_{00}-\bar{\mathbf{b}}^{\mathrm{T}}\mathbf{M}\bar{\mathbf{b}}\right]}da=$$

$$=\left[\begin{array}{l} t=\dfrac{a^2}{2}[(\mathbf{b}-\bar{\mathbf{b}})^{\mathrm{T}}\mathbf{M}(\mathbf{b}-\bar{\mathbf{b}})+m_{00}-\bar{\mathbf{b}}^{\mathrm{T}}\mathbf{M}\bar{\mathbf{b}}] \\ a=\sqrt{\dfrac{2t}{(\mathbf{b}-\bar{\mathbf{b}})^{\mathrm{T}}\mathbf{M}(\mathbf{b}-\bar{\mathbf{b}})+m_{00}-\bar{\mathbf{b}}^{\mathrm{T}}\mathbf{M}\bar{\mathbf{b}}}} \\ da=\dfrac{dt}{a[(\mathbf{b}-\bar{\mathbf{b}})^{\mathrm{T}}\mathbf{M}(\mathbf{b}-\bar{\mathbf{b}})+m_{00}-\bar{\mathbf{b}}^{\mathrm{T}}\mathbf{M}\bar{\mathbf{b}}]} \end{array}\right]=$$

$$=\frac{c2^{(N-n-2)/2}}{[(\mathbf{b}-\bar{\mathbf{b}})^{\mathrm{T}}\mathbf{M}(\mathbf{b}-\bar{\mathbf{b}})+m_{00}-\bar{\mathbf{b}}^{\mathrm{T}}\mathbf{M}\bar{\mathbf{b}}]^{(N-n)/2}}\int_0^\infty t^{(N-n-2)/2}e^{-t}dt,$$

čo využitím gama funkcie dáva

$$\pi_1(\mathbf{b}\,|\,\mathbf{x}) = \frac{c2^{(N-n-2)/2}\Gamma(\frac{N-n}{2})}{[(\mathbf{b}-\bar{\mathbf{b}})^{\mathrm{T}}\mathbf{M}(\mathbf{b}-\bar{\mathbf{b}})+m_{00}-\bar{\mathbf{b}}^{\mathrm{T}}\mathbf{M}\bar{\mathbf{b}}]^{(N-n)/2}} =$$

$$= \frac{c2^{(N-n-2)/2}\Gamma(\frac{N-n}{2})}{(m_{00}-\bar{\mathbf{b}}^{\mathrm{T}}\mathbf{M}\bar{\mathbf{b}})^{(N-n)/2}} \frac{(m_{00}-\bar{\mathbf{b}}^{\mathrm{T}}\mathbf{M}\bar{\mathbf{b}})^{(N-n)/2}}{[(\mathbf{b}-\bar{\mathbf{b}})^{\mathrm{T}}\mathbf{M}(\mathbf{b}-\bar{\mathbf{b}})+m_{00}-\bar{\mathbf{b}}^{\mathrm{T}}\mathbf{M}\bar{\mathbf{b}}]^{(N-n)/2}} =$$

$$c_1\left[\frac{(\mathbf{b}-\bar{\mathbf{b}})^{\mathrm{T}}\mathbf{M}(\mathbf{b}-\bar{\mathbf{b}})+m_{00}-\bar{\mathbf{b}}^{\mathrm{T}}\mathbf{M}\bar{\mathbf{b}}}{m_{00}-\bar{\mathbf{b}}^{\mathrm{T}}\mathbf{M}\bar{\mathbf{b}}}\right]^{-(N-n)/2} = c_1\left[1+\frac{(\mathbf{b}-\bar{\mathbf{b}})^{\mathrm{T}}\mathbf{M}(\mathbf{b}-\bar{\mathbf{b}})}{m_{00}-\bar{\mathbf{b}}^{\mathrm{T}}\mathbf{M}\bar{\mathbf{b}}}\right]^{-(N-n)/2},$$

kde

$$c_1 = \frac{c2^{(N-n-2)/2}\Gamma(\frac{N-n}{2})}{(m_{00}-\bar{\mathbf{b}}^{\mathrm{T}}\mathbf{M}\bar{\mathbf{b}})^{(N-n)/2}}.$$

□

Modus aposteriórneho rozdelenia $(\bar{a},\bar{\mathbf{b}})$ z vety 4.2., ktorý sme definovali v definícii 1.8., má zásadný význam, pretože sa využíva ako odhad parametrov $(a,\mathbf{b})$.

4.2 Odhady parametrov MA a ARMA modelu

Rozoberme najprv prípad MA modelu zadaného vzorcom (2.1) s fixovaným počiatkom ako u modelu AR. Zo spomenutého vzorca si vyjadríme Y_i ako

$$Y_i = a^{-1}X_i + \sum_{j=1}^{m} b_j Y_{i-j},\ i = m+1,\ ...,\ N, \tag{4.9}$$

kde $a = c_0$ a $b_j = \frac{c_j}{c_0}$ pre $j = 1,\ ...,\ m.$ Predpokladajme, že náhodné veličiny $Y_{n+1},\ ...,\ Y_N$ majú rozdelenie $\mathrm{N}(0,1)$ a rovnako ako v predošlom prípade, nech náhodná veličina a má apriórne rozdelenie a^{-1} pre $a > 0$ a náhodný vektor $\mathbf{b}$ apriórne rozdelenie rovné hodnote 1 v $\mathbb{R}^m$ nezávisle na $x_1,\ ...,\ x_m$. Združené aposteriórne rozdelenie náhodnej veličiny a a náhodného vektoru $\mathbf{b}$ nájdeme pomocou tvrdenia 1.6. Máme

$$g_i^{-1}(\mathbf{X}) = a^{-1}X_i + \sum_{j=1}^{m} b_j Y_{i-j}$$

a

$$\det \frac{\partial g^{-1}(\mathbf{x})}{\partial(\mathbf{x})} = \det \begin{vmatrix} a^{-1} & 0 & \cdots & 0 \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & a^{-1} \end{vmatrix} = a^{-(N-m)},$$

preto platí

$$r(x_{m+1},...,x_N \mid x_1,...,x_m,a,\mathbf{b}) = (2\pi)^{\frac{-(N-m)}{2}} a^{-(N-m)} e^{-\frac{1}{2a^2}\left[\sum_{j=m+1}^{N}(x_j-\sum_{i=1}^{m} b_i y_{j-i})^2\right]}.$$

Využitím vety 3.1. a poznámky 3.2. získavame združenú aposteriórnu hustotu náhodnej veličiny a a vektoru $\mathbf{b}$ v tvare

$$\pi(a,\mathbf{b} \mid \mathbf{x}) = c(2\pi)^{\frac{-(N-m)}{2}} a^{-(N-m+1)} e^{-\frac{1}{2a^2}\left[\sum_{m}^{N}(x_j-\sum_{i=1}^{m} b_i y_{j-i})^2\right]},$$

kde c je normujúca konštanta. Marginálnu aposteriórnu hustotu $\pi(\mathbf{b} \mid \mathbf{x})$ náhodného vektoru $\mathbf{b}$ získame ako

$$\int_0^\infty \pi(a,\mathbf{b} \mid \mathbf{x}) da = \left[\begin{array}{l} t = \frac{1}{2a^2}\left[\sum_{j=m+1}^{N}(x_j - \sum_{i=1}^{m} b_i y_{j-i})^2\right] \\ a = \sqrt{\frac{1}{2t}\left[\sum_{j=m+1}^{N}(x_j - \sum_{i=1}^{m} b_i y_{j-i})^2\right]} \\ da = \frac{-a^3 dt}{\left[\sum_{j=m+1}^{N}(x_j - \sum_{i=1}^{m} b_i y_{j-i})^2\right]} \end{array} \right] =$$

$$= c2^{(N-m-2)/2}(2\pi)^{-(N-m)/2}\left[\sum_{j=m+1}^{N}(x_j - \sum_{i=1}^{m} b_i y_{j-i})^2\right]^{-(N-m)/2} \int_0^\infty t^{(N-m-2)/2} e^{-t} dt.$$

Využitím gama funkcie a vytvorením konštanty c_1 získavame

$$\pi(\mathbf{b} \mid \mathbf{x}) = c_1\left[\sum_{j=m+1}^{N}(x_j - \sum_{i=1}^{m} b_i y_{j-i})^2\right]^{-(N-m)/2}.$$

Vidíme, že táto aposteriórna hustota, rovnako ako aj hustota $\pi(a,\mathbf{b} \mid \mathbf{x})$, je závislá na veličinách Y_1, ..., Y_{N-1}. Tie však nepoznáme a sú závislé práve na parametroch, ktoré sa snažíme odhadnúť a tak využitie týchto aposteriórnych hustôt pri odhadovaní parametrov nemá uplatnenie.

Teraz uvažujme model ARMA, ktorý je zadaný vzťahom 2.5. Nech má tento model takisto fixovaný počiatok. Úpravou zo vzťahu 2.5 dostávame

$$Y_i = a^{-1}(-\sum_{j=1}^{m} c_j Y_{i-j} + \sum_{k=0}^{n} d_k X_{i-k}),$$

kde $a = c_0$. Voľme apriórnu hustotu náhodnej veličiny a ako a^{-1} pre $a > 0$ a apriórne hustoty náhodných vektorov $\mathbf{c}$ a $\mathbf{d}$ rovné hodnote 1 v $\mathbb{R}^m$ a v $\mathbb{R}^{n+1}$. Teraz sa využije tvrdenie 1.6. a podobne ako v modeli MA povedie aposteriórna hustota

$\pi(a,\mathbf{c},\mathbf{d}\,|\,\mathbf{x})$ na nezápornú funkciu $g(a,\mathbf{c},\mathbf{d};X_{n+1},...,X_N;Y_1,...,Y_{N-1})$, v ktorej vystupujú práve ony neznáme náhodné veličiny Y_1, ..., Y_{N-1}. Z tohto dôvodu rovnako ako v modeli MA nemá táto cesta odhadovania parametrov využitie.

Na získanie odhadov sa využívajú takisto takzvané numerické metódy a jednou z nich je metóda ***Markov chain Monte Carlo*** (viď. pramene [3], [6]). Jej využitím v bayesovskej analýze, dokážeme na základe simulácií odhadovať parametre zložitých modelov akým model ARMA určite je. Avšak vzhľadom na určený rozsah práce tieto postupy odhadov parametrov uvádzať nebudeme.

5 Testy významnosti parametrov

V poslednej kapitole tejto práce, konkrétne v prvej podkapitole, sa budeme venovať základom teórie testovania parametrov. Uvedené poznatky sme čerpali zo zdroja [5]. V druhej podkapitole predstavíme kľúčovú vetu v testovaní významnosti parametrov modelu AR z prameňa [2] a zhrnutím predošlých vedomostí uvedieme tieto testy v poslednej podkapitole práce.

5.1 Úvod do teórie testovania

Testom významnosti parametra sa snažíme určiť, či je parameter daného modelu štatistický významný alebo štatisticky nevýznamný. Nástrojom určenia významnosti parametra je testovanie hypotézy o jeho nulovosti, preto musíme uviesť niekoľko základných pojmov a definícií z tejto oblasti štatistiky.

Majme náhodný výber $\mathbf{X} = (\mathbf{X}_1, \ldots, \mathbf{X}_n)^{\mathrm{T}}$ s rozdelením F, ktoré závisí na parametri $\boldsymbol{\theta} = (\theta_1, \ldots, \theta_m)^{\mathrm{T}}$, kde $\boldsymbol{\theta}$ patrí do parametrického priestoru Θ obsahujúceho aspoň dva rôzne prvky. Nech O predstavuje neprázdnu vlastnú podmnožinu priestoru Θ. Štatistickou hypotézou rozumieme tvrdenie, ktoré hovorí, že $\boldsymbol{\theta} \in \mathrm{O}$. Pomocou náhodného výberu $\mathbf{X}$ sa snažíme rozhodnúť o platnosti hypotézy a to pomocou vhodne zvolenej testovej štatistiky $S(\mathbf{X})$ a kritického oboru C. Rozdeľme teraz parametrický priestor Θ na dve disjunktné podmnožiny Θ_0 a Θ_1. Nulovou hypotézou rozumieme tvrdenie $\boldsymbol{\theta} \in \Theta_0$ s označením H_0 a alternatívnou hypotézou H_1 je tvrdenie $\boldsymbol{\theta} \in \Theta_1$.

V prípade, že naša vhodne zvolená funkcia dát $S(\mathbf{X})$, ktorá sa nazýva testová štatistika, spĺňa $S(\mathbf{X}) \in C$, hypotézu H_0 zamietame v prospech alternatívy H_1 a v opačnom prípade získavame záver, že hypotézu H_0 nemôžeme zamietnuť v prospech alternatívy H_1.

Na testovú štatistiku kladieme zásadnú požiadavku, aby jej rozdelenie bolo čo najviac citlivé na hodnotu testovaného parametra a čo najmenej závislé na ostatných charakteristikách rozdelenia s distribučnou funkciou F. Kladie sa preto podmienka, že pokiaľ $F_1 \neq F_2$ a parametre $\boldsymbol{\theta}_1$ a $\boldsymbol{\theta}_2$, na ktorých sú tieto rozdelenia závislé, sú zhodné, tak potom pre každú borelovskú množinu B má platiť

$$\lim_{n\to\infty}\left(\int \delta_B(S(\mathbf{x}))dF_1(x_1)...dF_1(x_n) - \int \delta_B(S(\mathbf{x}))dF_2(x_1)...dF_2(x_n)\right) = 0,$$

kde funkcia δ_B nadobúda hodnotu 1, pokiaľ $S(\mathbf{x}) \in B$ a inak hodnotu 0.

Požadujeme teda, aby rozdelenie testovej štatistiky bolo asymptoticky rovnaké, či už je náhodný výber $\mathbf{X}$ z rozdelenia F_1 alebo F_2. V prípade platnosti tejto podmienky, testová štatistika nezávisí okrem testovaného parametra $\boldsymbol{\theta}$ na žiadnych iných charakteristikách.

Definícia 5.1. Nech je dané $\alpha \in (0,1)$. Pokiaľ kritický obor C splňuje podmienku

$$\sup_{\boldsymbol{\theta}\in\Theta_0} \mathrm{P}_{\boldsymbol{\theta}}[S(\mathbf{X}) \in C] = \alpha,$$

povieme, že štatistický test $(S(\mathbf{X}), C)$ má hladinu α.

Poznámka 5.1. Hladinu testu chápeme ako pravdepodobnosť zamietnutia platnej hypotézy. Je štandardné voliť $\alpha = 0,05$.

Definícia 5.2. Interval $A \subset \mathbb{R}$ sa nazýva intervalový odhad parametra $\boldsymbol{\theta} \in \mathbb{R}$ so spoľahlivosťou $(1-\alpha)$, práve keď pravdepodobnosť pokrytia skutočnej hodnoty parametra $\boldsymbol{\theta}$ intervalom A je $(1-\alpha)$.

Ako sme sa už vyššie zmienili, na vyšetrovanie štatistickej významnosti parametra sa využívajú testovania hypotézy o jeho nulovosti. Konkrétnejšie sa dá povedať, že parameter je štatisticky významný, pokiaľ je zamietnutá nulová hypotéza o jeho nulovosti. Štatisticky nevýznamný je v prípade, že nulovú hypotézu nemôžeme zamietnuť, a preto je daný parameter natoľko nerozoznateľný od nuly, že jeho vynechanie z modelu, by nemalo viesť k veľkým zmenám. Tým sa náš model zjednoduší.

5.2 Nástroje k testovaniu významnosti parametrov

Uvažujme model AR s fixovaným počiatkom popísaným v úvode kapitoly 4. Nasledujúca veta nám ukáže spôsoby ako testovať parametre tohto modelu bayesovským prístupom.

Veta 5.1. Urobme predpoklad $\mathbf{M}^* > 0$ a $N > 2n,$ kde matice $\mathbf{M}^*$ a $\mathbf{M}$ predstavujú matice vo vete 4.2. Pokiaľ majú náhodné veličiny $a, \mathbf{b}$ aposteriórnu hustotu $\pi(a, \mathbf{b} \mid \mathbf{x})$ danú vzorcom (4.4), potom platí:

1) aposteriórne rozdelenie náhodnej veličiny

$$F = \frac{(N-2n)(\mathbf{b}-\bar{\mathbf{b}})^{\mathrm{T}}\mathbf{M}(\mathbf{b}-\bar{\mathbf{b}})}{n(m_{00}-\bar{\mathbf{b}}^{\mathrm{T}}\mathbf{M}\bar{\mathbf{b}})}$$

je F-rozdelenie o n a $N-2n$ stupňoch voľnosti,

2) vynechajme i-ty riadok a i-ty stĺpec matice $\mathbf{M}$ a takto vzniknutú maticu označme $\mathbf{M}_{ii}$ s determinantom $|\mathbf{M}_{ii}|$. Pre $n=i=1$ zadefinujme $|\mathbf{M}_{ii}|=1$. Potom aposteriórne rozdelenie každej veličiny danej ako

$$t_i = (b_i - \bar{b}_i)\left[\frac{|\mathbf{M}|(N-2n)}{|\mathbf{M}_{ii}|(m_{00}-\bar{\mathbf{b}}^{\mathrm{T}}\mathbf{M}\bar{\mathbf{b}})}\right]^{1/2},\ i=1,2,\ldots,n,$$

je Studentovo t rozdelenie s $N-2n$ stupňami voľnosti,

3) aposteriórne rozdelenie náhodnej veličiny

$$a^2(m_{00}-\bar{\mathbf{b}}^{\mathrm{T}}\mathbf{M}\bar{\mathbf{b}})$$

je χ^2 rozdelenie s $N-2n$ stupňami voľnosti.

Dôkaz: 1) Budeme vychádzať z toho, že náhodný vektor $\mathbf{b}$ má aposteriórnu hustotu π_1 danú vzorcom (4.7). Zaveďme vektor $\mathbf{z}=(z_1,\ \ldots,\ z_n)^{\mathrm{T}}$ rovnosťou

$$\mathbf{z} = (m_{00}-\bar{\mathbf{b}}^{\mathrm{T}}\mathbf{M}\bar{\mathbf{b}})^{-1/2}\mathbf{M}^{1/2}(\mathbf{b}-\bar{\mathbf{b}}),$$

z čoho vyplýva, že

$$\mathbf{b} = \mathbf{M}^{-1/2}(m_{00}-\bar{\mathbf{b}}^{\mathrm{T}}\mathbf{M}\bar{\mathbf{b}})^{1/2}\mathbf{z}+\bar{\mathbf{b}}. \tag{5.1}$$

Označme

$$\mathbf{N} = \mathbf{M}^{-1/2}(m_{00}-\bar{\mathbf{b}}^{\mathrm{T}}\mathbf{M}\bar{\mathbf{b}})^{1/2}.$$

Našim cieľom je nájsť aposteriórnu hustotu náhodného vektoru $\mathbf{z}$, a preto využijeme tvrdenie 1.6. Platí

$$g^{-1}(\mathbf{z}) = \mathbf{b} = \mathbf{M}^{-1/2}(m_{00}-\bar{\mathbf{b}}^{\mathrm{T}}\mathbf{M}\bar{\mathbf{b}})^{1/2}\mathbf{z}+\bar{\mathbf{b}} = \mathbf{N}\mathbf{z}+\bar{\mathbf{b}},$$

potom

$$g^{-1}(z_i) = \sum_{j=1}^{n}\mathbf{N}_{i,j}z_j + \bar{b}_i.\ \text{pre}\ i=1,\ \ldots,\ n,$$

kde $\mathbf{N}_{i,j}$ sú prvky vzniknutej matice $\mathbf{N}$. A tak $\det\dfrac{\partial g^{-1}(\mathbf{z})}{\partial \mathbf{z}}$ je výraz, ktorý nezávisí na $\mathbf{z}$ a môžeme ho zahrnúť do novej konštanty c_2. Po dosadení vzťahu (5.1) do vzorca (4.7), celkovo získavame

$$\pi_2(\mathbf{z}\,|\,\mathbf{x}) = c_2\left[1+\mathbf{z}^{\mathrm{T}}\mathbf{z}\right]^{-(N-n)/2}.$$

Ak sa pozrieme na náhodnú veličinu F, je zrejmé, že

$$F = \frac{N-2n}{n}\mathbf{z}^{\mathrm{T}}\mathbf{z} = \frac{N-2n}{n}\sum\nolimits_{i=1}^{n} z_i^2 .$$

Uvažujme teraz náhodné veličiny $y,\ \beta_1,\ ...,\ \beta_{n-1}$ pre, ktoré platí, že

$$y \geq 0,\ \ 0 \leq \beta_1,\ ...,\ \beta_{n-2} < \pi,\ \ 0 \leq \beta_{n-1} < 2\pi.$$

Urobíme ďalšiu transformáciu vektorov a opäť využijeme tvrdenie 1.6. Nech teda

$$\begin{aligned}
z_1 &= y^{1/2}\cos\beta_1,\\
z_2 &= y^{1/2}\sin\beta_1\cos\beta_2,\\
&.\\
&.\\
z_{n-1} &= y^{1/2}\sin\beta_1 ... \sin\beta_{n-2}\cos\beta_{n-1},\\
z_n &= y^{1/2}\sin\beta_1 ... \sin\beta_{n-2}\sin\beta_{n-1}.
\end{aligned}$$

Označme $\mathbf{w} = (y, \beta_1, ..., \beta_{n-1})^{\mathrm{T}}$. Platí

$$\det\frac{\partial g^{-1}(\mathbf{w})}{\partial\mathbf{w}} = \frac{1}{2}y^{(n/2)-1}h(\beta_1, ..., \beta_{n-1}),$$

kde funkcia h nijako nezávisí na y a vzhľadom na podmienky, ktoré kladieme na y a β_i, je nezáporná. Celkovo získavame združenú aposteriórnu hustotu náhodných veličín $y,\ \beta_1,\ ...,\ \beta_{n-1}$ ako

$$\pi_3(y, \beta_1, ..., \beta_{n-1} \mid \mathbf{x}) = c_3 y^{(n/2)-1}\ \ 1+y\ ^{-(N-n)/2}\ h(\beta_1, ..., \beta_{n-1}),$$

keďže $\mathbf{z}^{\mathrm{T}}\mathbf{z} = \sum\nolimits_{i=1}^{n} z_i^2 = y.$ Teraz nájdeme marginálnu aposteriórnu hustotu $y.$ Zintegrujeme veličiny $\beta_1,\ ...,\ \beta_{n-1}$ cez obory, na ktorých sme si ich definovali, čím získame konštantu, ktorá nezávisí na $y.$ Zahrňme ju do konštanty $c_4,$ preto marginálna hustotu y je v podobe

$$\pi_4(y \mid \mathbf{x}) = c_4 y^{(n/2)-1}\ \ 1+y\ ^{-(N-n)/2} \quad \text{pre } y > 0.$$

Nakoniec urobme transformáciu $F = \dfrac{N-2n}{n}y.$ A tak $g^{-1}(F) = \dfrac{n}{N-2n}F$ a $\det\dfrac{\partial g^{-1}(F)}{\partial} = \dfrac{n}{N-2n}.$ Získavame aposteriórnu hustotu náhodnej veličiny F ako

$$\pi_5(f \mid \mathbf{x}) = c_4\frac{n}{N-2n}\left(\frac{n}{N-2n}f\right)^{(n/2)-1}\left[1+\frac{n}{N-2n}f\right]^{-(N-n)/2} =$$

$$= c_5(f)^{(n/2)-1}\left[1+\frac{n}{N-2n}f\right]^{-(N-n)/2} \quad \text{pre } f > 0.$$

A to je hustota F-rozdelenia s n a $N-2n$ stupňami voľnosti. Hoci nie je priamo vidieť, že konštanty c_5 a konštanta z rozdelenia F zavedeného v podkapitole 1.2. sú zhodné, tak ich rovnosť nutne platí, pretože integrály oboch hustôt na obore $(0,\infty)$ nadobúdajú hodnotu 1.

2) Nech $i=1$. Z poznámky 3.3. vieme, že symetrickú pozitívne definitnú maticu $\mathbf{M}$ je možné vyjadriť v tvare $\mathbf{M}=\mathbf{C}^{\mathrm{T}}\mathbf{C}$, kde $\mathbf{C}=(c_{ij})_{i,j=1}^{N}$ je dolná trojuholníková matica, a preto

$$\mathbf{C}\mathbf{M}^{-1}\mathbf{C}^{\mathrm{T}}=\mathbf{I}. \tag{5.2}$$

Zadefinujme si teraz maticu $\mathbf{D}$ vzťahom

$$\mathbf{M}^{-1}=\mathbf{D}=(d_{ij})_{i,j=1}^{N}.$$

Uvažujme vektor $\mathbf{z}=(z_1,\ ...,\ z_n)^{\mathrm{T}}$ a nech $\mathbf{z}=\mathbf{C}(\mathbf{b}-\bar{\mathbf{b}})$. Keďže $\mathbf{C}$ je dolná trojuholníková matica, tak jednak z (5.2) vyplýva, že $c_{11}^2 d_{11}=1$ a zároveň potom, že

$$z_1=d_{11}^{-1/2}(b_1-\bar{b}_1).$$

Pre prvok d_{11} platí vzťah

$$d_{11}=|\mathbf{M}_{11}|\,/\,|\mathbf{M}| \text{ pre } n\ge 1, \tag{5.3}$$

kde matica $\mathbf{M}_{11}$ vznikla vynechaním prvého stĺpca a riadku a pre $n=i=1$ sme zadefinovali $|\mathbf{M}_{ii}|=1$. Z predpokladu $\mathbf{M}^*>0$ vo vete 4.2 vyplýva, že i $d_{11}>0$. Celkovo teda

$$(\mathbf{b}-\bar{\mathbf{b}})^{\mathrm{T}}\mathbf{M}(\mathbf{b}-\bar{\mathbf{b}})=\mathbf{z}^{\mathrm{T}}\mathbf{z}=\frac{(b_1-\bar{b}_1)^2}{m_{11}}+\sum\nolimits_{j=2}^{n} z_j^2.$$

Transformujme $(b_1,b_2...,b_n)$ na $(b_1,z_2,...,z_n)$. Potom $\det\dfrac{\partial g^{-1}(b_1,z_2,...,z_n)^{\mathrm{T}}}{\partial(b_1,z_2,...,z_n)^{\mathrm{T}}}$ je konštanta, ktorá nezávisí na $(b_1,z_2,...,z_n)$ a môžeme ju zahrnúť do konštanty c_6, ktorá vznikla z c. Celkovo máme

$$\pi_6(b_1,z_2,...,z_n|\mathbf{x})=c_6\left[1+\frac{(b_1-\bar{b}_1)^2/m_{11}+\sum_{j=2}^{n} z_j^2}{m_{00}-\bar{\mathbf{b}}^{\mathrm{T}}\mathbf{M}\bar{\mathbf{b}}}\right]^{-(N-n)/2}.$$

Teraz vypočítame marginálnu aposteriórnu hustotu parametra b_1. Máme

$$\pi_7(b_1|\mathbf{x})=\int_{\mathbb{R}}...\int_{\mathbb{R}}\pi(b_1,z_2,...,z_n|\mathbf{x})dz_2...dz_n=c_7\left[1+\frac{(b_1-\bar{b}_1)^2/m_{11}}{m_{00}-\bar{\mathbf{b}}^{\mathrm{T}}\mathbf{M}\bar{\mathbf{b}}}\right]^{-(N-2n+1)/2}.$$

Uskutočnením transformácie

$$t_1 = (b_1 - \bar{b}_1)\left[\frac{|\mathbf{M}|(N-2n)}{|\mathbf{M}_{11}|(m_{00} - \bar{\mathbf{b}}^{\mathrm{T}}\mathbf{M}\bar{\mathbf{b}})}\right]^{1/2}$$

a využitím (5.3) získavame aposteriórnu hustotu náhodnej veličiny t_1 v tvare

$$\pi_8(t|\mathbf{x}) = c_8\left[1 + \frac{t^2}{N-2n}\right]^{-(N-2n+1)/2},$$

ktorý zodpovedá Studentovému rozdeleniu s $N-2n$ stupňami voľnosti. Prípad, v ktorom $i > 1,$ sa dokazuje obdobne.

3) Pripomeňme si aposteriórnu hustotu veličín $a, \mathbf{b}$ danú vzorcom 4.4. Našim cieľom je teraz nájsť marginálnu aposteriórnu hustotu veličiny a. Teda

$$\pi_9(a \mid \mathbf{x}) = ca^{N-n-1}e^{\frac{-a^2}{2}[m_{00} - \bar{\mathbf{b}}^{\mathrm{T}}\mathbf{M}\bar{\mathbf{b}}]}\int_{\mathbb{R}}\ldots\int_{\mathbb{R}} e^{\frac{-a^2}{2}[(\mathbf{b}-\bar{\mathbf{b}})^{\mathrm{T}}\mathbf{M}(\mathbf{b}-\bar{\mathbf{b}})}db_1\ldots db_n,$$

čo následne využitím gama funkcií dá aposteriórnu hustotu veličiny a v tvare

$$\pi_9(a \mid \mathbf{x}) = c_9 a^{N-2n-1}e^{\frac{-a^2}{2}[m_{00} - \bar{\mathbf{b}}^{\mathrm{T}}\mathbf{M}\bar{\mathbf{b}}]},\ a > 0.$$

Transformáciou $y = a^2(m_{00} - \bar{\mathbf{b}}^{\mathrm{T}}\mathbf{M}\bar{\mathbf{b}})$ získame aposteriórnu hustotu veličiny y ako

$$\pi_{10}(y \mid \mathbf{x}) = c_{10}y^{\frac{(N-2n)}{2}-1}e^{\frac{-y}{2}},\ y > 0.$$

Tá zodpovedá rozdeleniu χ^2 s $N-2n$ stupňami voľnosti.

□

5.3 Uvedenie testov významnosti

Prvý bod vety 5.1. nám poskytuje možnosť ako testovať hypotézu, že skutočná hodnota parametra $\mathbf{b}$ je $\mathbf{b}^{\bullet}$. Do F dosadíme za $\mathbf{b}$ vektor $\mathbf{b}^{\bullet}$ a v prípade, ak vypočítaná hodnota F prekročí kritickú hodnotu F-rozdelenia s n a $N-2n$ stupňami voľnosti, potom túto hypotézu zamietame. Pri testoch významnosti parametra $\mathbf{b}$ volíme $\mathbf{b}^{\bullet} = \mathbf{0},$ a tým zároveň testujeme, že $\{X_t\}$ je postupnosť nekorelovaných náhodných veličín proti alternatíve, že tvorí autoregresný model radu nanajvýš n. Náš test významnosti vyzerá nasledovne.

Testovaný parameter:

$$\mathbf{b} = (b_1,\ b_2,\ \ldots,\ b_n)^{\mathrm{T}}.$$

Hypotéza a alternatíva:

$$\mathrm{H}_0 : \mathbf{b} = \mathbf{0},\ \mathrm{H}_1 : \mathbf{b} \neq \mathbf{0}.$$

Testová štatistika:

$$F = \frac{(N-2n)(\mathbf{b}-\bar{\mathbf{b}})^{\mathrm{T}}\mathbf{M}(\mathbf{b}-\bar{\mathbf{b}})}{n(m_{00}-\bar{\mathbf{b}}^{\mathrm{T}}\mathbf{M}\bar{\mathbf{b}})}.$$

Kritický obor:

$$\mathrm{H}_0 \text{ zamietame} \Leftrightarrow F \leq F_{n,N-2n}(\alpha/2) \vee F \geq F_{n,N-2n}(1-\alpha/2).$$

$F_{n,N-2n}(\alpha/2)$ a $F_{n,N-2n}(1-\alpha/2)$ predstavujú po poradí $(\alpha/2)$-ty a $(1-\alpha/2)$-ty kvantil F-rozdelenia s n a $N-2n$ stupňami voľnosti. Hodnota α vyjadruje hladinu testu podľa definície 5.1.

Druhý bod vety 5.1. uplatníme v prípade, že testujeme štatistickú významnosť konkrétnej zložky b_i parametra $\mathbf{b}$. V prípade testovania $\mathrm{H}_0 : b_n = 0,$ sa snažíme zistiť, či je možné, aby model AR, ktorý popisuje postupnosť $\{X_t\},$ bol radu $n-1$ namiesto n.

Testovaný parameter:

$$b_i,\ i = 1, ..., n.$$

Hypotéza a alternatíva:

$$\mathrm{H}_0 : b_i = 0,\ \mathrm{H}_1 : b_i \neq 0.$$

Testová štatistika:

$$t_i = (b_i - \bar{b}_i)\left[\frac{|\mathbf{M}|(N-2n)}{|\mathbf{M}_{ii}|(m_{00}-\bar{\mathbf{b}}^{\mathrm{T}}\mathbf{M}\bar{\mathbf{b}})}\right]^{1/2}.$$

Kritický obor:

$$\mathrm{H}_0 \text{ zamietame} \Leftrightarrow |t_i| \geq t_{N-2n}(1-\alpha/2).$$

Výraz $t_{N-2n}(1-\alpha/2)$ reprezentuje $(1-\alpha/2)$-ty kvantil t-rozdelenia s $N-2n$ stupňami voľnosti.

A nakoniec tretí bod vety 5.1. sa predovšetkým uplatňuje pri stanovovaní intervalu spoľahlivosti pre a^2. Pokúsime sa teraz skonštruovať tento interval. Z definície 5.3. vieme, že hľadáme interval, pre ktorý platí

$$P[A \ni a^2] = 1-\alpha.$$

Označme $\chi^2_{N-2n}(\alpha)$ ako α -kvantil χ^2 rozdelenia s $N-2n$ stupňami voľnosti, potom

$$P[\chi^2_{N-2n}(\alpha/2) < a^2(m_{00} - \bar{\mathbf{b}}^{\mathrm{T}}\mathbf{M}\bar{\mathbf{b}}) < \chi^2_{N-2n}(1-\alpha/2)] = 1-\alpha.$$

Keďže podľa nerovnice (4.8) je výraz $(m_{00} - \bar{\mathbf{b}}^{\mathrm{T}}\mathbf{M}\bar{\mathbf{b}})$ kladný, tak úpravami získavame

$$P\left[\frac{\chi^2_{N-2n}(\alpha/2)}{(m_{00} - \bar{\mathbf{b}}^{\mathrm{T}}\mathbf{M}\bar{\mathbf{b}})} < a^2 < \frac{\chi^2_{N-2n}(1-\alpha/2)}{(m_{00} - \bar{\mathbf{b}}^{\mathrm{T}}\mathbf{M}\bar{\mathbf{b}})}\right] = 1-\alpha.$$

Tým sme našli interval *A*, ktorý s pravdepodobnosťou $(1-\alpha)$ pokrýva druhú mocninu parametra a.

6 Záver

Práca ma teoretický charakter. V jej kontexte sme sa venovali teórii pravdepodobnosti a štatistiky a základom z oblasti náhodných procesov. Ďalej sme uviedli model ARMA a modely, ktoré ho spoločne vytvárajú. V neposlednom rade sme si predstavili bayesovskú analýzu a konjugované systémy apriórnych hustôt a ukázali sme možnosti voľby týchto apriórnych hustôt. Súčasne sme tieto voľby využili pri hľadaní aposteriórnych hustôt v pravdepodobnostných rozdeleniach uvedených v práci. Prístup, ktorý bayesovská analýza ponúka, sme prvotne využili v odhadoch parametrov kľúčových modelov tejto práce a následne v testoch významnosti parametrov modelu AR. Práve tieto časti práce považujeme za najdôležitejšie, i keď je potrebné dodať, že významnou kapitolou je aj samotný pohľad a preniknutie do sveta bayesovskej analýzy.

Parametre, ktoré v modeloch vystupujú, výrazne časové rady ovplyvňujú, a preto je ich charakterizovanie vo finančných modeloch z hľadiska určenia budúceho vývoja ekonomických veličín zásadné. Uplatnenie práce tak vidíme hlavne v modelovaní náhodných postupností AR, MA a ARMA a to určením odhadov, záverov a testov významnosti o ich parametroch. Poznatky práce môžeme aplikovať pri modelovaní náhodných postupností a to nielen v ekonomike a vo finančnom sektore, ale rovnako tak aj v oblasti životného prostredia, v medicíne, a všade tam, kde sa časové rady vyskytujú.

Literatúra

[1] ANDĚL, J.: 1978. *Matematická statistika*. Praha: SNTL/ALFA, 1978.352 s.

[2] ANDĚL J.: 1976. *Statistická analýza časových řád*. Praha: STNL, 1976. 272 s.

[3] HAMILTON, J. D.: 1994. *Time Series Analysis*. Princenton, New Jersey: Princeton University Press, 1994. 820 s. ISBN 0-691-04289-6.

[4] HUŠKOVÁ M.:1985. *Bayesovské metody*. Praha: Univerzita Karlova, 1985. 96 s.

[5] KULICH, M.: 2012. *Malý větníček.*[online].[cit.2012-03-13]. Dostupné na: <http://www.karlin.mff.cuni.cz/~kulich/>.

[6] MONAHAN, J. F.: *Fully Bayesian analysis of ARMA times series models*. In: Journal of Econometrics, 1983, č. 21, s. 307-330.

[7] PRÁŠKOVÁ Z.: 2004. *Základy náhodných procesů II* . Praha: Karolinum, 2004. 152 s. ISBN 80-246-0871-1.